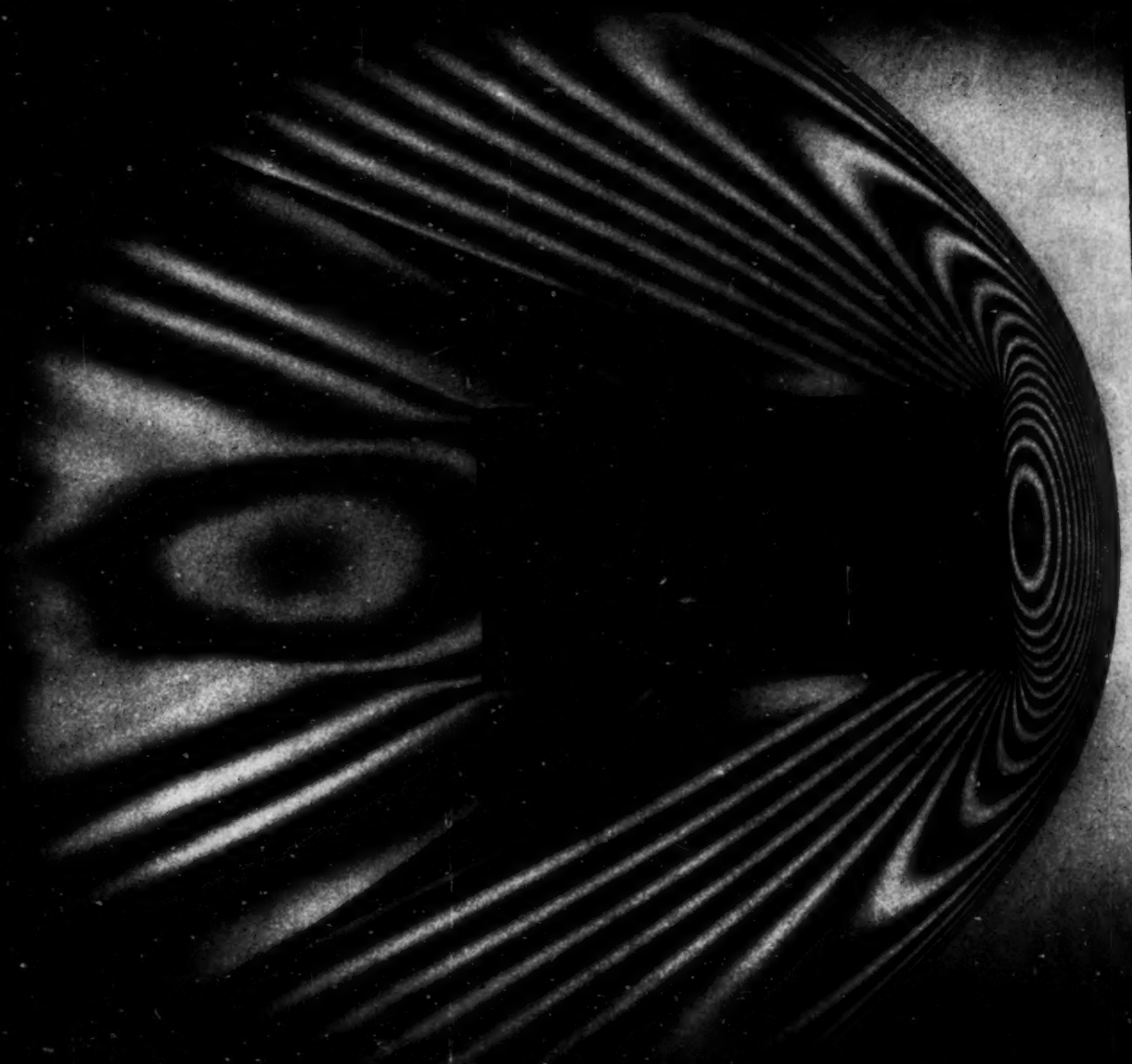
JOURNAL OF APPLIED PHYSICS



JOURNAL
OF
APPLIED PHYSICS

Vol. 23, No. 8 AUGUST, 1952

In This Issue

Page

X-Ray Diffraction by Face-Centered Cubic Crystals with Deformation Faults M. S. Paterson 805

Some Limitations on the Maximum Frequency of Coherent Oscillations Robert S. Elliott 812

Mathematical Analysis of a Dielectric Amplifier Louis A. Pipes 818

Emission of Radiation from Diatomic Gases. IV. Emissivity Calculations for CO and HCl for Nonoverlapping Rotational Lines as a Function of Temperature and Optical Density S. S. Penner 825

Time-Temperature Dependence of Linear Viscoelastic Behavior F. Schwarzl and A. J. Staverman 838

Estimate of the Energies of the Positive Ions in a Negative Point-to-Plane Corona G. L. Weissler and Mark Schindler 844

On the Theory of Antenna Beam Shaping Allen S. Dunbar 847

Point Focusing X-Ray Monochromator for the Study of Low Angle Diffraction Leon Shenfil, Warren E. Danielson, and Jesse W. M. DuMond 854

Latex Particle Size Determination Using Diffraction Peaks Obtained with the Point Focusing X-Ray Monochromator Warren E. Danielson, Leon Shenfil, and Jesse W. M. DuMond 860

Anelasticity of Zinc Shuichiro Takahashi 866

Decay and Recovery of the Pulsed Emission of Oxide-Coated Cathodes R. M. Matheson and L. S. Nergaard 869

Measurement of the Internal Energy in Copper Introduced by Cold Work Benjamin Welber 876

Production of Monochromatic X-Radiation for Microradiography by Excitation of Fluorescent Characteristic Radiation T. H. Rogers 881

Longitudinal and Transverse Ultrasonic Waves in a Synthetic Rubber A. W. Nolle and P. W Sieck 888

Space Charge Reduction of Low Frequency Fluctuations in Thermionic Emitters T. B. Tomlinson 894

Electron Microscopy of Synthetic Latices Samuel H. Maron, Carl Moore, and Allen S. Powell 900

Dirac Delta-Function and the Summation of Fourier Series M. R. Spiegel 906

Electromagnetic Back-Scattering from Cylindrical Wires C. T. Tai 909

Thermal Force on an Aerosol Particle in a Temperature Gradient R. L. Saxton and W. E. Ranz 917

Free Molecular Flow in the Sample Inlet to the Mass Spectrometer Paul D. Zemany 924

Integrable Case of Electron Motion in Electric and Magnetic Field H. Poritsky and R. P. Jerrard 928

Letters to the Editor:

Possible Explanation of Briggs' Limiting Negative Pressure Data V. J. Berry and E. W. Hough 931

Temperature Gradient Owing to Radiation Absorption G. B. Shook 931

Single Synthetic Cadmium Sulfide Crystals S. J. Czyzak, D. J. Craig, and C. E. McCain 932

Announcements 934

Index to Advertisers xxx

BOARD OF EDITORS
Elmer Hutchisson,
Editor
Earle C. Gregg,
Assistant Editor

Associate Editors
Howard H. Aiken
Ray E. Bolz
Lyle D. Borst
Robert D Heidenreich
James Hillier
Leonard B. Loeb
W. James Lyons
Eric Reissner
Roman Smoluchowski

Publication of the
American Institute of Physics

OFFICERS
George R. Harrison,
Chairman
George B. Pegram,
Treasurer

GOVERNING BOARD
Karl K. Darrow
G. J. Dienes
S. A. Goudsmit
G. P. Harnwell
George R. Harrison
F. V. Hunt
Rudolph Kingslake
Hugh S. Knowles
C. C. Lauritsen
F. W. Loomis
William F. Meggers
George B. Pegram
I. I. Rabi
Duane Roller
Frederick Seitz
Francis G. Slack
John C. Steinberg
Richard M. Sutton
Mark W. Zemansky

Consultant on Printing
Melvin Loos

ADMINISTRATIVE STAFF
Henry A. Barton,
Director
Wallace Waterfall,
Executive Secretary
Theodore Vorburger,
Advertising Manager
Ruth F. Bryans,
Publication Manager
Betty H. Goodfriend,
Assistant Secretary
Eileen B. Neuberger,
Circulation Manager
Kathryn Setze
Assistant Treasurer

The Journal of Applied Physics, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as Physics. The Journal of Applied Physics publishes editorials and reviews, as well as technical papers of applied physics. Articles appearing in it are indexed regularly, according to their field, by Physics Abstracts, Chemical Abstracts, Engineering Index, Inc., Industrial Arts Index, etc.

Manuscripts should be submitted to Elmer Hutchisson, Editor, Case Institute of Technology, Cleveland 6, Ohio. Unless otherwise stated, submission of a manuscript is a representation that it has been neither copyrighted, published, nor submitted for publication elsewhere. The authors' institutions are requested to pay a **publication charge** of $8 per page which, if honored, entitles them to 100 reprints without further charge. Instructions will be sent with galley proofs.

Subscriptions, renewals, and orders for back numbers should be addressed to the American Institute of Physics, 57 East 55 Street, New York 22, New York.

Subscription Price	U.S. and Canada	Elsewhere
To members of the American Institute of Physics and Affiliated Societies	$10.00	$11.00
To all others	12.00	13.00

Back Number Prices

Yearly back number rate when complete year is available: $14.00. Single copies: $1.25 each.

Changes of address should be sent to the American Institute of Physics; **corrected proofs** should be sent to the Publication Manager, American Institute of Physics. **Advertising rates** supplied on request. Orders, advertising copy, and cuts should be sent to the American Institute of Physics.

Entered as second class matter January 22, 1937, at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1879. Accepted for mailing at the special rate of postage provided for in paragraph (d-2), section 34.40, P. L. & R. of 1948, authorized May 2, 1932.

Journal
of
Applied Physics

Volume 23, Number 8 August, 1952

X-Ray Diffraction by Face-Centered Cubic Crystals with Deformation Faults

M. S. Paterson
*Institute for the Study of Metals, University of Chicago, Chicago, Illinois**
(Received December 31, 1951)

In view of the recent interest in the introduction of stacking faults in crystals during plastic deformation, the x-ray diffraction effects have been calculated for face-centered cubic crystals with deformation faults (as distinct from growth faults, the expressions for which are given in an appendix for comparison). Explicit expressions are derived for the reciprocal lattice intensity distribution, and it is shown how the intensity of faulting can be determined from measurements on the breadths of the reciprocal lattice streaks or the displacements of their peaks. It is then shown how these effects will be revealed in the x-ray powder pattern and some practical aspects are discussed.

1. INTRODUCTION

WHEN a crystal lattice is conceived in terms of the stacking of layers in a well-defined manner, any error in the regular sequence of the layers is known as a stacking "fault," "mistake," or "disorder." From x-ray diffraction studies it is known that stacking faults may arise in the growth of crystals with well-developed layer structures, for example, micas,[1,2] graphite,[3] silicon carbide,[3] uranyl chloride,[4] and others. Prominent faulting can also arise during martensitic phase transformations, for example, in cobalt[5] and lithium.[6]

More recently, Barrett[7] has demonstrated that stacking faults can be introduced in crystals during plastic deformation; this may be of fundamental importance in connection with a number of aspects of the plastic deformation and microstructure of metals. Thus, Barrett[8] suggests that faulting may sometimes contribute significantly to the x-ray line broadening that results from plastic deformation. A detailed knowledge of the x-ray diffraction effects to be expected from structures faulted during deformation should, therefore, be of value in future studies. This paper presents the results of a calculation for face-centered cubic crystals in which the faults occur in (111) planes; the detailed intensity distribution in the reciprocal lattice, the positions of the intensity maxima and their breadths, and other results are expressed in terms of the density of faulting.

At this point, the distinction between the faulted structure arising in growth and that resulting from plastic deformation should be made clear. The ideal face-centered cubic structure can be described simply as an $A\,B\,C\,A\,B\,C\cdots$ stacking sequence of close-packed (111) planes, where the atoms in the B and C layers are situated above the alternative sets of hollows between the atoms in the A layers; a fault is an error in this regular sequence, with the restriction that adjacent layers must always be different. Thus, the growth of a crystal by the addition of successive (111) layers is governed by the rule that each new layer is different from the last two except when a fault occurs; then the new layer is identical with the second last layer. $A\,B\,C\,A\,B\,(A)\,C\,B\,A\,C\,B\,A\cdots$ is such a sequence with a fault at the sixth layer. On the other hand, in the case of faulting in deformation, a regular packing sequence is supposed already to exist before a fault is introduced, so the packing sequence is always the same except at the fault plane; the rule of packing is that the original sequence is always followed except where a fault occurs,

* This work was done during leave of absence from Aeronautical Research Laboratories, Department of Supply, Melbourne, Australia.

[1] S. B. Hendricks and M. E. Jefferson, Am. Mineral. **24**, 729 (1939); S. B. Hendricks, Phys. Rev. **57**, 448 (1940).
[2] S. B. Hendricks and E. Teller, J. Chem. Phys. **10**, 147 (1942).
[3] H. Jagodzinski and F. Laves, Schweiz. mineralog. petrog. Mitt. **28**, 456 (1948); H. Jagodzinski, Acta Cryst. **2**, 298 (1949).
[4] W. H. Zachariasen, Acta Cryst. **1**, 277 (1948).
[5] O. S. Edwards and H. Lipson, Proc. Roy. Soc. (London) **180A**, 268 (1942).
[6] C. S. Barrett, Trans. Am. Inst. Mining Met. Engrs. **175**, 605 (1948). (Reply to discussion on Barrett and Trautz, p. 579.)
[7] C. S. Barrett, Trans. Am. Inst. Mining Met. Engrs. **188**, 123 (1950).
[8] C. S. Barrett, Phys. Rev. **81**, 311 (1951).

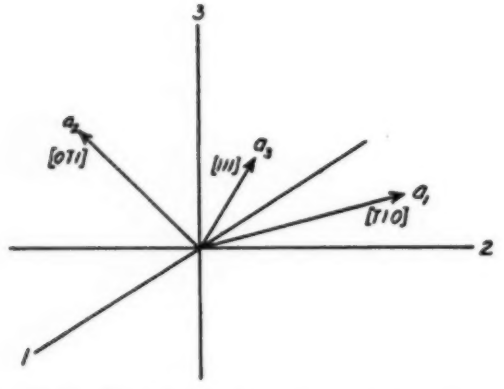


FIG. 1. Relationship between the chosen hexagonal axes $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ and the cubic axes 1, 2, 3.

thus $A\,B\,C\,A\,B\,(A)\,B\,C\,A\,B\,C\cdots$. We see, therefore, that in the growth case the fault plane is the boundary between two portions of crystal having a twin orientational relationship, whereas in a crystal faulted by deformation all regions between the fault planes are of parallel orientation. Alternatively, a deformation fault could be described as the juxtaposition of two growth faults (or "twin faults").

The theory of x-ray diffraction by a hexagonal close-packed lattice with growth faulting has been given by Wilson;[9] this can also be applied to growth faulting in face-centered cubic crystals (Appendix II). Zachariasen[10] and Méring[11] have given general treatments of diffraction by faulted lattices, and Barrett[7] has applied Zachariasen's theory to deformation faulting in the face-centered cubic lattice. However, the present paper treats the problem more fully, giving explicit results and showing several ways in which comparison with experimental observations can be made.

2. CALCULATION OF THE INTENSITY DISTRIBUTION IN RECIPROCAL SPACE

(a) Preliminary

This calculation is based to some extent on Wilson's treatment of faulting in hexagonal cobalt,[9] but the approach is also similar to that of Warren and Averbach[12] to x-ray line broadening in cold-worked metals. Following Wilson[9] and Barrett,[7] it is convenient to choose hexagonal axes in the crystal, the third axis being normal to the faulting plane (111) and the other two axes being parallel to it, with directions $[\bar{1}10]$ and $[0\bar{1}1]$ relative to the cubic axes (see Fig. 1, where the cubic axes are 1, 2, 3 and the hexagonal axes are indicated by the unit vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$, whose lengths relative to the cubic unit cell edge are $1/\sqrt{2}$, $1/\sqrt{2}$, and $\sqrt{3}$). Thus we shall use hexagonal indices $(HK\cdot L)$ referred to this unit cell of three atoms.

[9] A. J. C. Wilson, Proc. Roy. Soc. (London) **180A**, 277 (1942).
[10] W. H. Zachariasen, Phys. Rev. **71**, 715 (1947).
[11] J. Méring, Acta Cryst. **2**, 371 (1949).
[12] B. E. Warren and B. L. Averbach, J. Appl. Phys. **20**, 885 (1949); **21**, 595 (1950).

If the unit vector $\mathbf{s}_0$ defines the direction of incidence of an x-ray beam on the crystal, the intensity of scattering in the direction of a unit vector $\mathbf{s}$ is given by

$$I=\sum_j\sum_{j'} F_jF_{j'}^*\exp\left[\frac{2\pi i}{\lambda}(\mathbf{s}-\mathbf{s}_0)\cdot(\mathbf{r}_j-\mathbf{r}_{j'})\right],$$

where F_j, $F_{j'}$ are the structure amplitudes of the unit cells at the positions defined by the radius vectors $\mathbf{r}_j$, $\mathbf{r}_{j'}$; each summation is over all unit cells in the crystal.

Putting

$$\mathbf{s}-\mathbf{s}_0/\lambda=h_1\mathbf{b}_1+h_2\mathbf{b}_2+h_3\mathbf{b}_3,$$

where $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ are the hexagonal reciprocal lattice base vectors and h_1, h_2, h_3 are continuous variables,

$$\mathbf{r}_j=j_1\mathbf{a}_1+j_2\mathbf{a}_2+j_3\mathbf{a}_3$$

and

$$j_1'=j_1+m_1,$$

etc. (the m may be positive or negative), we have

$$I=\sum_j\sum_m F_jF_{j+m}^*\exp\left[2\pi i\left(h_1m_1+h_2m_2+\frac{h_3m_3}{3}\right)\right].$$

For convenience, we have here taken F_j to be the scattering amplitude, with appropriate phase, for a single atom; the factor $\frac{1}{3}$ appears in the third term of the exponent because h_3 refers to a unit cell of three atoms in height (see Wilson[9]).

We can now write

$$I=\sum_m V_mJ_m\exp\left[2\pi i\left(h_1m_1+h_2m_2+\frac{h_3m_3}{3}\right)\right], \quad (1)$$

where $J_m=\langle F_jF_{j+m}^*\rangle$ is the mean value of the product $F_jF_{j+m}^*$ for all pairs of cells $m_1\mathbf{a}_1+m_2\mathbf{a}_2+m_3\mathbf{a}_3$ apart, and V_m is the fraction of the cells for which the end of this translation vector lies within the crystal. This equation is equivalent to that given by Wilson.[4]

We shall assume for simplicity that the smallness of size of the crystal is not important (that is, that J_m is already small before V_m departs appreciably from unity as $|m_1\mathbf{a}_1+m_2\mathbf{a}_2+m_3\mathbf{a}_3|$ increases), so V_m can be taken as unity. Also, F_j and F_{j+m} can be thought of as the scattering amplitudes for atoms belonging to close-packed layers $m_3\mathbf{a}_3$ apart and so J_m can be computed in terms of the density of faults and the value of m_3 since there is no disturbance in the regular arrangement within the close-packed layers. That is, $J_m=J_{m_3}$, independent of m_1 and m_2. Thus we have from (1)

$$I(h_1h_2h_3)=\sum_{m_1}\exp[2\pi ih_1m_1]\cdot\sum_{m_2}\exp[2\pi ih_2m_2]\cdot\sum_{m_3}J_{m_3}\exp\left[2\pi i\frac{h_3m_3}{3}\right],$$

which is only nonzero when $h_1=H$, $h_2=K$, H and K

being integers, giving

$$I(HKh_3)=C\sum_{m_3}J_{m_3}\exp\left[2\pi i\frac{h_3m_3}{3}\right], \qquad (2)$$

where C is a constant.

This expresses the fact that any broadening must correspond to streaks in the $\mathbf{b}_3$ direction in the reciprocal lattice (normal to the plane of faulting), there being no significant broadening in the $\mathbf{b}_1$ or $\mathbf{b}_2$ directions unless the crystal is very small.

(b) Calculation of J_m†

Since the quantities F refer to single atoms, there are three possible values of F corresponding to A, B, or C layers and differing from each other only in phase due to the relative lateral displacements of layers in the regular stacking sequence. Thus, taking the scattering power of single atoms as unity, the three values of F are

$$F_a=1, \quad F_b=\exp\left[\frac{2\pi i}{\lambda}\left(\frac{\mathbf{a}_1+2\mathbf{a}_2}{3}\right)\cdot(\mathbf{s}-\mathbf{s}_0)\right],$$

$$F_c=\exp\left[\frac{2\pi i}{\lambda}\left(\frac{2\mathbf{a}_1+\mathbf{a}_2}{3}\right)\cdot(\mathbf{s}-\mathbf{s}_0)\right].$$

We now introduce $P_m{}^0$, $P_m{}^+$, and $P_m{}^-$ to represent the probabilities that two layers m apart are in an $A\cdots A$, $A\cdots B$, or $A\cdots C$ relationship, respectively:

$$P_m{}^0+P_m{}^++P_m{}^-=1.$$

Then

$$\begin{aligned}J_m=1/3\{&F_aF_a{}^*P_m{}^0+F_aF_b{}^*P_m{}^++F_aF_c{}^*P_m{}^-\\&+F_bF_b{}^*P_m{}^0+F_bF_c{}^*P_m{}^++F_bF_a{}^*P_m{}^-\\&+F_cF_c{}^*P_m{}^0+F_cF_a{}^*P_m{}^++F_cF_b{}^*P_m{}^-\}.\end{aligned}$$

After inserting the F values, this reduces to

$$J_m=P_m{}^0+P_m{}^+\exp\left[2\pi i\frac{H-K}{3}\right]+P_m{}^-\exp\left[-2\pi i\frac{H-K}{3}\right]. \qquad (3)$$

If we denote by α the probability that a deformation fault will occur at any given layer in an $A\,B\,C\,A\,B\,C\cdots$ crystal, it can be shown (Appendix I) that, for positive m,

$$\left.\begin{aligned}P_m{}^0&=1/3+2/3[-(1-3\alpha(1-\alpha))^{\frac{1}{2}}]^m\cos m\theta\\P_m{}^+&=1/3-1/3[-(1-3\alpha(1-\alpha))^{\frac{1}{2}}]^m\\&\quad\times(\cos m\theta+\sqrt{3}\sin m\theta)\\P_m{}^-&=1/3-1/3[-(1-3\alpha(1-\alpha))^{\frac{1}{2}}]^m\\&\quad\times(\cos m\theta-\sqrt{3}\sin m\theta)\end{aligned}\right\}, \qquad (4)$$

where $\theta=\arctan\sqrt{3}(1-2\alpha)$. Also

$$P_{-m}{}^0=P_m{}^0; \quad P_{-m}{}^+=P_m{}^-; \quad P_{-m}{}^-=P_m{}^+.$$

It may be noted that J_m of Eq. (3) is identical with Zachariasen's quantity $W_M{}^{H_1H_2}$, used by Barrett,[7] but we now have explicit expressions for the P_m.

(c) Final Expressions for the Intensity Distribution

For $H-K=3N$, N any integer, it is seen from Eq. (3) that always $J_m=1$ and so Eq. (2) gives sharp intensity peaks for $h_3=L=3N'$, N' any integer, and zero intensity elsewhere. Thus we have sharp maxima at $L=3N'$ for all values of α, that is, for both perfect and faulted crystals.

For $H-K=3N\pm1$, it is convenient to write Eq. (2) in the form

$$I(HKh_3)=C\sum_{n=0}^{\infty}\left(a_n\cos 2\pi n\frac{h_3}{3}+b_n\sin 2\pi n\frac{h_3}{3}\right).$$

From Eqs. (3) and (4) it can then be shown that

$$\begin{aligned}&a_0=1; \quad a_n=2[-(1-3\alpha(1-\alpha))^{\frac{1}{2}}]^n\cos n\theta \quad (n\neq0)\\&b_0=0; \quad b_n=\pm2[-(1-3\alpha(1-\alpha))^{\frac{1}{2}}]^n\sin n\theta \quad (n\neq0)\end{aligned} \qquad (5)$$

(the $\pm$ sign refers to $H-K=3N\pm1$, respectively). Then by further rearrangement we have

$$I(HKh_3)=C\left\{1+2\sum_1^{\infty}[(1-3\alpha(1-\alpha))^{\frac{1}{2}}]^n\times\cos 2\pi n\left(\frac{h_3}{3}+\tfrac{1}{2}\pm\frac{\theta}{2\pi}\right)\right\}, \qquad (6)$$

which can be summed ($\alpha\neq0$, 1) to give

$$I(HKh_3)=C\frac{\frac{3}{2}\alpha(1-\alpha)}{1-\frac{3}{2}\alpha(1-\alpha)-[1-3\alpha(1-\alpha)]^{\frac{1}{2}}\cos 2\pi n[(h_3/3)+\frac{1}{2}\pm(\theta/2\pi)]}. \qquad (7)$$

This is a symmetrical intensity distribution with its peak occurring where $(h_3/3)+\frac{1}{2}\pm(\theta/2\pi)=N'$, N' any integer, i.e.,

$$h_3=3N'-\frac{3}{2}\mp\frac{3}{2\pi}\arctan\sqrt{3}(1-2\alpha). \qquad (8)$$

Its integral breadth, defined as

$$\beta=\frac{\int_{\text{peak }h_3-3/2}^{\text{peak }h_3+3/2}I(HKh_3)dh_3}{I_{\max}(HKh_3)},$$

† The suffix 3 in m_3 will be omitted in the remainder of this paper.

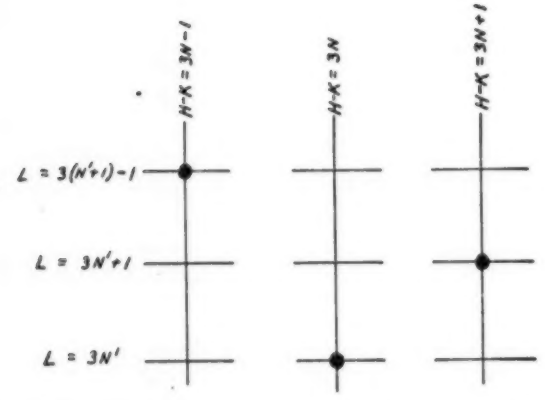


FIG. 2. Reciprocal lattice representation of the x-ray reflections from a perfect face-centered cubic crystal.

can be calculated from (6) or (7) to be

$$\beta=3\cdot\frac{1-[1-3\alpha(1-\alpha)]^{\frac{1}{2}}}{1+[1-3\alpha(1-\alpha)]^{\frac{1}{2}}}. \tag{9}$$

When $\alpha=0$ it is seen from Eq. (6) that sharp intensity peaks occur at the positions

$$h_3=L=3N'\pm1.$$

Similarly when $\alpha=1$, sharp intensity peaks occur at the positions

$$h_3=L=3N'\mp1.$$

These are the maxima for the perfect crystal which is in a twin orientation with respect to the chosen axes.

3. DESCRIPTION OF THE RECIPROCAL LATTICE

The intensity distribution of diffraction can be conveniently represented in the following way. We think of the reciprocal lattice in terms of rows of points parallel to the $\mathbf{b}_3$ directions, each row having H, K indices belonging to one of the three classes $H-K=3N$, $3N\pm1$ and each point in a given row being distinguished by $L=3N'$, $3N\pm1$. Then all the intensity distributions can be described with reference to a diagram such as Fig. 2. The black dots in Fig. 2 represent the positions of the sharp reflections from a perfect face-centered cubic crystal.

Figure 3 represents how, when faults are introduced in the crystal with increasing density, the reflections with $H-K=3N$ and $L=3N'$ remain sharp but the other reflections broaden and their maxima shift in opposite directions. The shaded areas indicate an intensity distribution along the $H-K=3N\pm1$ rows, there being zero intensity of scattering for non-integral values of h_1 and h_2. The shape of the intensity distribution of these streaks remains symmetrical, the exact form being given by (6) or (7) and the integral breadth by (9), while the positions of the maxima are given by (8), all in terms of the probability α of a fault occurring at any layer.

For $\alpha=\frac{1}{2}$, the maxima of the broadened reflections all occur at values of $h_3=3N'+3/2$ and the integral breadth has its greatest value, 1. As α increases further, the reflections gradually decrease in breadth and approach more nearly the sharp reflections for the perfect crystal of twin orientation $A\,C\,B\,A\,C\,B\,\cdots$, indicated by the open circles in Fig. 3.

If the crystal were sufficiently small for ordinary diffraction broadening to occur (the term V_m in (1) must then be retained to express the effect), this would appear as a three-dimensional blurring of all the reflections represented in Fig. 3, including $H-K=3N$, in accordance with the shape transform of the crystal. All reflections would be influenced in the same way by the small size effect.

A summary of the diffraction effects in a *growth-faulted crystal* is given in Appendix II in a form that permits direct comparison with the reciprocal lattice pattern for a deformation-faulted crystal. The essential difference is revealed in the asymmetry of the components of the double-peaked reflections for the lesser degrees of faulting and in the continuous transition to the diffraction pattern for the hexagonal close-packed structure as α approaches unity.

4. THE POWDER DIFFRACTION PATTERN

(a) General

Since polycrystalline or powdered samples with random orientation are often used for convenience in x-ray studies, we shall now discuss briefly how the reciprocal lattice intensity distribution for a deformation-faulted crystal will be revealed in the x-ray powder pattern. Each powder pattern line is formed by the superposition of the reflections with a given Bragg angle (that is, a given radial distance from the center of the reciprocal lattice), for all possible orientations. For example, the $\{\bar{1}11\}$ line is made up of the (111), (11$\bar{1}$), (1$\bar{1}$1), ($\bar{1}$11), ($\bar{1}$1$\bar{1}$), ($\bar{1}\bar{1}$1), ($\bar{1}\bar{1}\bar{1}$), and (1$\bar{1}\bar{1}$) reflections.

With the aid of the relations between the cubic

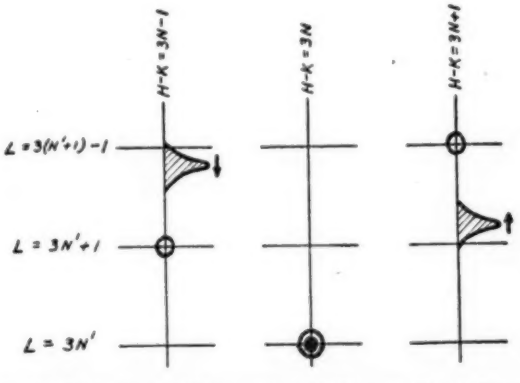


FIG. 3. Reciprocal lattice representation of the influence of deformation faulting on the x-ray reflections from a face-centered cubic crystal.

FIG. 4. The influence of deformation faulting on the x-ray powder pattern; the heavy vertical lines represent the sharp components and the shaded areas the diffuse components, while the indices indicate the positions of the reflections from perfect material.

indices (hkl) and the hexagonal indices $(HK\cdot L)$,

$$h=\frac{-2(2H+K)+L}{3},\quad k=\frac{2(H-K)+L}{3},$$

$$l=\frac{2(H+2K)+L}{3},$$

we can find the $(HK\cdot L)$ indices for each reflection that contributes to a given $\{hkl\}$ line, and then, by determining whether $L=3N$, $3N\pm1$, we can decide which components are sharp and which are broadened, and in which direction from the Bragg angle each broadened component is displaced (the displacement in the powder pattern is towards higher or lower angles depending on whether the displacement in the reciprocal lattice is away from or towards the plane $L=0$, respectively). However, since $L=h+k+l$ and since $-L=3N\mp1$ is equivalent to $L=3N\pm1$ in representing displacement away from or towards the plane $L=0$, the effect of each component on the powder pattern can be determined directly from the value of $|h+k+l|$ without explicit reference to the hexagonal indices. Thus, if N is any integer, $|h+k+l|=3N$ gives a sharp reflection, and $|h+k+l|=3N\pm1$ gives a broadened reflection displaced towards higher or lower Bragg angles, respectively. For example, (111) and $(\bar{1}\bar{1}\bar{1})$ give sharp components ($|h+k+l|=3$), and the others of the {111} family give broadened components ($|h+k+l|=1$) displaced to a higher angle.

By considering in this way each of the $\{hkl\}$ families that give the powder pattern lines, the effect of deformation faulting on the whole pattern can be deduced. It is shown diagrammatically in Fig. 4, where the heavy vertical lines represent the sharp components and the shaded areas the diffuse components, the arrows indicating the direction of displacement of the latter as α increases. The numbers in Fig. 4 are the multiplicity factors for the components. It should be noted, however, that Fig. 4 applies only to the ideal case in which each crystal contains only one parallel set of faults.

The quantitative details of the powder pattern are readily obtained from Eqs. (7), (8), and (9) by using the appropriate multiplying factor K for converting from the reciprocal lattice to the powder pattern intensity distribution (for example, see Warren and Averbach[12]); the resultant line profile is obtained by summing the contributions from the various components, sharp and diffuse. The intensity distribution, the displacements of the peaks and the integral breadths of the broadened components are completely determined in terms of α; conversely, from measurements of the peak displacements or integral breadths of the broadened components, α can be obtained (other means of obtaining α are mentioned later).

(b) Effect on Warren and Averbach Analysis

Warren and Averbach[12] have shown that if the intensity distribution across a powder diffraction line is expressed as a Fourier series, $I(x)=C\sum A_n\cos nx$, vital information for distinguishing between different causes of broadening is revealed in the behavior of A_n with increasing n. The effect of fault broadening can be seen from Eq. (6), where $A_n=[(1-3\alpha(1-\alpha))^{\frac{1}{2}}]^n$; the curve of A_n *versus* n will have an initial tangent of $\log[1-3\alpha(1-\alpha)]^{\frac{1}{2}}$, finite for $\alpha\neq0$. Thus, if only powder lines consisting of one type of diffuse components are considered, the effect of faulting on the A_n versus n curve is qualitatively similar to that of small size; from the initial slope of the curve the value of α can be obtained.

(c) Practical Application

So far it has been assumed that the faults are distributed at random and are confined to one parallel set of planes in each crystal. However, from the microstructure of copper-silicon alloys (Barrett[7]) it seems likely that the faults are not distributed at random but are clustered on a microscopic scale. In this case, the expressions for the intensity distribution of the broadening will not be strictly true, although qualitatively the behavior should be similar. When extreme clustering occurs, the x-ray effects can be thought of as the superposition of those for heavy faulting and for very slight or no faulting.

In polycrystalline material, especially after heavy deformation, faults can also be expected on several nonparallel sets of planes in each crystal, in analogy with multiple slip. Then, as a first approximation, the effects of each parallel set of faults may be superposed in the reciprocal lattice, and all the (hkl) components of each powder line (Fig. 4) are likely to be broadened to some extent.

The x-ray study of deformation faulting is further complicated in practice by the presence of the usual line broadening that accompanies cold work. However, if

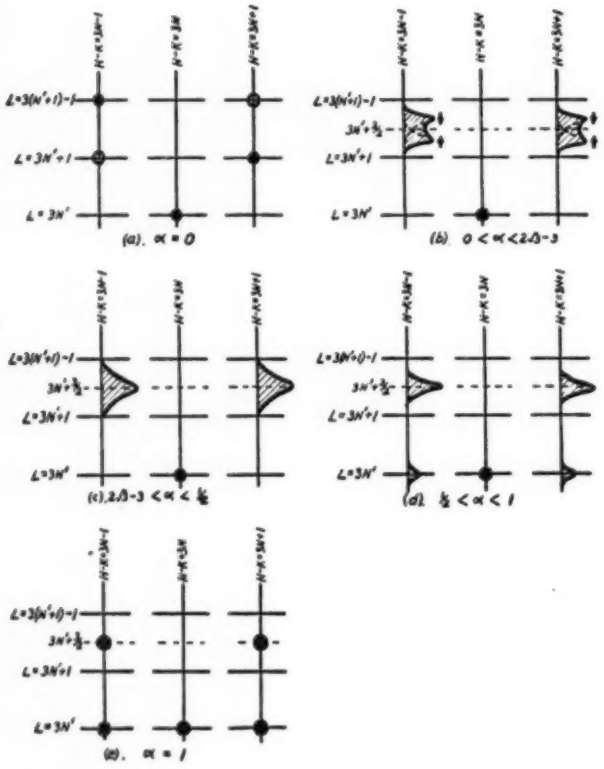


FIG. 5. Reciprocal lattice representation of the influence of growth faulting on the x-ray reflections from a face-centered cubic crystal, where α is the probability of a fault occurring at any (111) plane.

we assume that this broadening due to small size effects and internal strains is symmetrical about the Bragg angle (see Warren and Averbach[12]), a relation depending only on α can be established between the cosine and sine Fourier coefficients for the resulting intensity distribution when fault broadening is present. Thus if the broadening due to cold-work effects other than faults is expressed as $K\sum c_n \cos nx$, the fault broadening expressed as $KC\sum(a_n' \cos nx + b_n' \sin nx)$ where a_n' and b_n' are obtained from (5) by changing the origin from zero angle to the Bragg angle, and the resultant broadening expressed as $K\sum(\alpha_n \cos nx + \beta_n \sin nx)$, we have, from Stokes' relations[13] between the Fourier coefficients,

$$\frac{\beta_n}{\alpha_n}=\frac{b_n'}{a_n'}=\pm\tan n\left(\arctan\sqrt{3}(1-2\alpha)+\frac{2\pi}{3}\right). \quad (10)$$

(It is important to note that the Bragg angle must be chosen as origin in obtaining the Fourier coefficients α_n, β_n from the measured intensity distribution curve.) Relation (10) thus provides a practical means of determining α, subject to the limitations mentioned in the last two paragraphs. However, it is only valid when all (hkl) components of the powder line have the same intensity distribution and it must be modified if components of different intensity distributions are superposed.

[13] A. R. Stokes, Proc. Phys. Soc. (London) **61**, 382 (1948).

ACKNOWLEDGMENT

The author is grateful to Professor C. S. Barrett for stimulating discussions of this work.

APPENDIX I. DERIVATION OF THE P_m FOR DEFORMATION FAULTING

Only the calculation of $P_m{}^0$ will be given since $P_m{}^+$ and $P_m{}^-$ can be found in a similar way. The method is based on that used by Wilson.[9]

The zeroth layer can always be called A. We then represent the four possible arrangements leading to the mth layer being identical with it as

0	$m-2$	$m-1$	m
A	..... A	B	A
A	..... A	C	A
A	..... B	C	A
A	 C	B	A.

If α is the probability of a fault occurring at any layer in the $A\ B\ C\cdots$ crystal, $1-\alpha$ is the probability of A being followed by B, B by C, etc., while α is the probability of A being followed by C, etc. Thus the sum of the probabilities of each of the above arrangements is

$$P_m{}^0=P_{m-2}{}^0\cdot(1-\alpha)\alpha+P_{m-2}{}^0\cdot\alpha(1-\alpha)+P_{m-2}{}^+(1-\alpha)(1-\alpha)+P_{m-2}{}^-\cdot\alpha\cdot\alpha.$$

Also we have

$$P_{m-1}{}^0=P_{m-2}{}^+\cdot\alpha+P_{m-2}{}^-\cdot(1-\alpha)$$

and

$$P_{m-2}{}^0+P_{m-2}{}^++P_{m-2}{}^-=1.$$

From these we obtain the difference equation for $P_m{}^0$:

$$P_m{}^0+P_{m-1}{}^0+\{1-3\alpha(1-\alpha)\}P_{m-2}{}^0=1-\alpha(1-\alpha).$$

The solution of this, with $P_0{}^0=1$ and $P_1{}^0=0$, is

$$P_m{}^0=\tfrac{1}{3}+\tfrac{1}{3}\left\{-\tfrac{1}{2}+\frac{i\sqrt{3}}{2}(1-2\alpha)\right\}^m+\tfrac{1}{3}\left\{-\tfrac{1}{2}-\frac{i\sqrt{3}}{2}(1-2\alpha)\right\}^m,$$

which can be rearranged in the form

$$P_m{}^0=1/3+2/3\{-[1-3\alpha(1-\alpha)]^1\}^m\times\cos\{m\arctan\sqrt{3}(1-2\alpha)\}.$$

Similarly, by considering the possible arrangements of layers leading to B and C at the mth layer, we obtain the expressions for $P_m{}^+$ and $P_m{}^-$, respectively.

APPENDIX II. GROWTH FAULTING IN FACE-CENTERED CUBIC CRYSTALS

Since a face-centered cubic lattice is equivalent to a close-packed hexagonal lattice with a growth fault at every (00·1) layer, the treatment of x-ray diffraction by a face-centered cubic lattice with growth faulting on (111) planes follows directly from Wilson's treatment of growth faulting in hexagonal cobalt[9] when $(1-\alpha)$ is substituted for α. However, further calculation is required to obtain comparable results to those given for deformation faulting in this paper. The final expressions

and the description of the reciprocal lattice will not be given.

For all values of α, a sharp $(HK\cdot L)$ reflection occurs for $H-K=3N$ and $L=3N'$, N and N' being integers (Fig. 5).

For $H-K=3N\pm1$, the reflections become broadened in the $\mathbf{b}_3$ direction as α increases from zero (Fig. 5). Figure 5(a), for $\alpha=0$, shows the usual sharp f.c.c. reflections ($L=3N'\pm1$, respectively) as black dots, and also the reflections from the twin orientation (open circles) which also appear as soon as a single growth fault is introduced.

As α increases in the range $0<\alpha<2\sqrt{3}-3$, the $H-K=3N\pm1$ reflections broaden and their maxima move closer together (Fig. 5(b)); the intensity distribution can be expressed as

$$I(HKh_3)=C\Big\{1+\sum_1^\infty[(1-2\alpha)^{\frac{1}{2}}]^n\Big[\cos2\pi n\Big(\frac{h_3}{3}+\tfrac{1}{2}+\frac{\theta}{2\pi}\Big)+\frac{\alpha}{(3(1-2\alpha)-\alpha^2)^{\frac{1}{2}}}\sin2\pi n\Big(\frac{h_3}{3}+\tfrac{1}{2}+\frac{\theta}{2\pi}\Big)\Big]+\sum_1^\infty[(1-2\alpha)^{\frac{1}{2}}]^n\Big[\cos2\pi n\Big(\frac{h_3}{3}+\tfrac{1}{2}-\frac{\theta}{2\pi}\Big)-\frac{\alpha}{(3(1-2\alpha)-\alpha^2)^{\frac{1}{2}}}\sin2\pi n\Big(\frac{h_3}{3}+\tfrac{1}{2}-\frac{\theta}{2\pi}\Big)\Big]\Big\},$$

where

$$\theta=\arctan\frac{(3(1-2\alpha)-\alpha^2)^{\frac{1}{2}}}{1-\alpha};$$

the maxima occur at $h_3=3N'+(3/2)\pm(3\theta/2\pi)$ and both approach $h_3=3N'+3/2$ as $\alpha\to2\sqrt{3}-3$. Also it is seen that the two components have asymmetrical intensity distributions. Clearly, the resultant intensity distribution is different from that for the superposition of two deformation-faulted crystals of twin orientation.

For $2\sqrt{3}-3<\alpha<\frac{1}{2}$, the $H-K=3N\pm1$ reflections have an intensity distribution that is symmetrical about a maximum at $h_3=3N+3/2$ (Fig. 5(c)), and can be expressed as

$$I(HKh_3)=C\Big\{1+\sum_1^\infty\Big[\Big(1-\frac{\alpha}{[\alpha^2-3(1-2\alpha)]^{\frac{1}{2}}}\Big)\times\Big(\frac{1-\alpha}{2}-\tfrac{1}{2}(\alpha^2-3(1-2\alpha)^{\frac{1}{2}}\Big)^n+\Big(1+\frac{\alpha}{[\alpha^2-3(1-2\alpha]^{\frac{1}{2}}}\Big)\Big(\frac{1-\alpha}{2}+\tfrac{1}{2}(\alpha^2-3(1-2\alpha))^{\frac{1}{2}}\Big)^n\Big]\cos2\pi n\Big(\frac{h_3}{3}+\tfrac{1}{2}\Big)\Big\},$$

where both quantities with exponent n are positive and less than unity.

For $\frac{1}{2}<\alpha<1$, the intensity distribution for $H-K=3N\pm1$ (Fig. 5(d)) can be expressed as

$$I(HKh_3)=C\Big\{1+\Big(1-\frac{\alpha}{[\alpha^2-3(1-2\alpha)]^{\frac{1}{2}}}\Big)\times\sum_1^\infty\Big[-\frac{1-\alpha}{2}+\tfrac{1}{2}(\alpha^2-3(1-2\alpha))^{\frac{1}{2}}\Big]^n\cos2\pi n\frac{h_3}{3}+\Big(1+\frac{\alpha}{(\alpha^2-3(1-2\alpha))^{\frac{1}{2}}}\Big)\sum_1^\infty\Big[\frac{1-\alpha}{2}+\tfrac{1}{2}(\alpha^2-3(1-2\alpha))^{\frac{1}{2}}\Big]^n\cos2\pi n\Big(\frac{h_3}{3}+\tfrac{1}{2}\Big)\Big\},$$

where again both quantities with exponent n are positive and less than unity. This has maxima at $h_3=3N'$ and $3N'+3/2$. As $\alpha\to1$, the intensity distributions decrease in breadth until for $\alpha=1$ sharp reflections occur for $L=3N'$ and $3N'+3/2$; these are the reflections for the perfect close-packed hexagonal lattice (Fig. 5(e)).

Some Limitations on the Maximum Frequency of Coherent Oscillations*

ROBERT S. ELLIOTT†
Naval Ordnance Laboratory, White Oak, Maryland
(Received February 5, 1952)

Electron beam oscillators are placed in one of two groups depending on whether or not they employ resonant energy-extractors. A natural upper frequency limit exists for the group which does. This limit depends chiefly on the ac beam current density and the noise level in the resonant structure. For practical values of these parameters, the maximum frequency is below the range of visible light.

INTRODUCTION

IN recent years, experimenters have been concerned with the generation of coherent electromagnetic waves at higher and higher frequencies. The use of radar in World War II led to the development of oscillators operating at wavelengths as short as one centimeter. Postwar researches have endeavored to extend this range to the millimeter and sub-millimeter wavelengths. Speculation has even been raised as to whether or not it is possible to generate the visible spectrum coherently. However, there is general agreement on one point: the higher the frequency, the greater the difficulty.

The traditional method of achieving coherent oscillations has been to incorporate a resonant energy-extractor in the generating device. This resonant structure might be composed of a coil and condenser in parallel, a transmission line section, a cavity, etc. The distinguishing feature of all such structures is that they consist of conductors which are shaped for a very definite purpose. This purpose is to constrain the free electrons within the conductors to have a macroscopic motion which is periodic. The geometry of the structure dictates the periods of the motion and this gives rise to a set of natural frequencies of oscillation.

The electromagnetic field supported by the resonant structure consists essentially of a standing wave whenever one of the resonant frequencies is excited. This wave has an integral (or zero) number of space half-cycles in each of three independent directions. The three integers which specify this information are known as the mode numbers. The lowest resonant frequency corresponds to the mode with the fewest space half-cycles (called the fundamental mode). Which mode is excited depends on the manner in which energy is fed into the resonant structure.

It is usually simplest to excite the fundamental mode and most resonant structures are designed for such operation. Recognition of the fact that the size of a resonant structure oscillating in its fundamental mode is of the order of a free-space wavelength raises two serious technological limitations: (1) At short wavelengths, the required resonant structure is so small that fabrication becomes a problem; (2) if it is desired to extract even a reasonable amount of power from the resonant structure, the required field intensities are high enough to cause field emission with all its attendant troubles.

These difficulties can be avoided by designing resonant structures to operate in higher modes, thus achieving the same frequency with a physically larger structure. However, attempts to do this have not been too satisfactory. This suggests that perhaps the limitation is fundamental rather than technological.

It is the purpose of this paper to show that for certain oscillators the limitation is indeed fundamental. For convenience of discussion, oscillators will be divided into two groups. Group I contains all coherent oscillators which require a resonant structure to extract energy from an electron beam, the useful output being obtained by coupling to the field of the resonant structure. Group II contains all oscillators which do not fit the description of Group I.

An upper bound on the frequency for all Group I oscillators will be established. This bound depends on the values of the operating parameters. It is a necessary condition but not a sufficient one. Hence the least upper bound may be expected to be considerably lower. This will be evident from the assumptions made in the course of the derivation.

The analysis is applicable to klystrons and magnetrons and to all conventional vacuum tube oscillators. It is not applicable to traveling wave tubes, nor to oscillators in which the bunched electron beam radiates directly.

Rationalized *mks* units will be used throughout this paper.

THE THEOREM OF LIMITATION

The following general theorem serves to establish a frequency bound for all Group I oscillators.

For any coherent oscillator which functions by virtue of the interaction between a bunched electron beam and the electromagnetic field supported by a resonant structure, the following equation must be

* The research reported in this paper was supported in part by Contract N6-ori-71, Task XIX between the ONR and the University of Illinois.
† Formerly with the Department of Electrical Engineering, University of Illinois, Urbana, Illinois.

satisfied if a useful output is to be obtained: namely,

$$I_0^2 \geqslant \frac{2\omega^2\epsilon}{Q^2V} U^{Av}, \tag{1}$$

in which

I_0 is the "average" ac beam current density,
$\nu=(\omega/2\pi)$ is the resonant frequency,
ϵ is the permittivity of the medium,
Q is the loaded quality factor of the resonant structure,
V is the volume of the region in which the beam and the resonant field interact, and
U^{Av} is the average signal energy stored in the resonant field.

Proof. Consider any Group I oscillator. Let P_S^{Av} be the average power supplied by the beam to the field of the resonant energy-extractor of the oscillator. Let P_L^{Av} be the total average power lost by the resonant structure (including useful output and ohmic loss). If the oscillator is functioning,

$$P_S^{Av} = P_L^{Av}. \tag{2}$$

If $\mathbf{I}(x,y,z,t)$ is the current density function of the beam, and if $\mathbf{E}(x,y,z)\cdot\sin\omega t$ is the electric force function of the resonant field, (for which $\nu=(\omega/2\pi)$ is the resonant frequency), then

$$P_S^{Av} = \frac{1}{T}\int_0^T dt \int_V [\mathbf{I}(x,y,z,t)dS]\cdot[\mathbf{E}(x,y,z)\sin\omega t\cdot dl], \tag{3}$$

where $T=(2\pi/\omega)$ is the period of oscillation and the volume integration extends over the region of interaction of the beam and the resonant field. Further,

$$P_S^{Av} = \frac{1}{T}\int_0^T dt \int_V [\mathbf{I}'(x,y,z)\cdot\sin\{\omega t+\delta(x,y,z)\}\cdot dS] \cdot[\mathbf{E}(x,y,z)\sin\omega t\cdot dl], \tag{4}$$

in which $\mathbf{I}'$ is that portion of the function $\mathbf{I}$ which is periodic at the frequency $\nu=(\omega/2\pi)$. This is true because no other time function combines with the expression for the electric force to yield an average power different from zero.

It is thus apparent that

$$P_S^{Av} \leqslant \frac{1}{2}\int_V |\mathbf{I}'(x,y,z)|\cdot|\mathbf{E}(x,y,z)|\cdot dV, \tag{5}$$

because the largest value for the integral would occur if the current density were everywhere colinear with the electric force and in time phase with it.

Let I_0 be a constant defined by

$$\int_V I_0\cdot|\mathbf{E}(x,y,z)|\cdot dV = \int_V |\mathbf{I}'(x,y,z)|\cdot|\mathbf{E}(x,y,z)|\cdot dV. \tag{6}$$

If I_{max} and I_{min} are respectively the maximum and minimum values of $|\mathbf{I}'(x,y,z)|$ throughout the interaction volume V, then it is apparent that

$$I_{min} \leqslant I_0 \leqslant I_{max}. \tag{7}$$

Thus I_0 is seen to be a weighted average of the ac beam current density in the interaction region, the weighting function being the resonant electric field distribution. Equation (2) can now be rewritten

$$I_0^2/4\left\{\int_V |\mathbf{E}(x,y,z)|\cdot dV\right\}^2 \geqslant \{P_L^{Av}\}^2. \tag{8}$$

Define the loaded Q of the resonant structure in the usual way, that is,

$$Q=\omega U^{Av}/P_L^{Av}, \tag{9}$$

in which U^{Av} is the average signal energy stored in the resonant field. Equation (8) becomes

$$I_0^2/4\left\{\int_V |\mathbf{E}(x,y,z)|\cdot dV\right\}^2 \geqslant \{\omega U^{Av}/Q\}^2. \tag{10}$$

But, by Schwarz's inequality,[1]

$$\left\{\int_V dV\right\}\left\{\int_V |\mathbf{E}(x,y,z)|^2\cdot dV\right\} \geqslant \left\{\int_V |\mathbf{E}(x,y,z)|\cdot dV\right\}^2, \tag{11}$$

so that

$$I_0^2V/4\int_V |\mathbf{E}(x,y,z)|^2\cdot dV \geqslant \{\omega U^{Av}/Q\}^2. \tag{12}$$

If V_1 is the volume of the entire region occupied by the resonant field, then it is obvious that $V_1 \geqslant V$ and also

$$I_0^2V/4\int_{V_1} |\mathbf{E}(x,y,z)|^2\cdot dV \geqslant \{\omega U^{Av}/Q\}^2. \tag{13}$$

The electric stored energy is given by

$$U_E = \epsilon/2\int_{V_1} |\mathbf{E}(x,y,z)|^2\cdot\sin^2\omega t\cdot dV, \tag{14}$$

and at resonance U^{Av} equals twice the average electric stored energy. Thus,

$$U^{Av} = \epsilon/2\int_{V_1} |\mathbf{E}(x,y,z)|^2\cdot dV, \tag{15}$$

where ϵ is the permittivity of the medium. Finally, (13) becomes

$$I_0^2 \geqslant \frac{2\omega^2\epsilon}{Q^2V} U^{Av}. \tag{1}$$

[1] W. V. Lovitt, *Linear Integral Equations* (McGraw-Hill Book Company, Inc., New York), first edition, p. 125.

(1) may be used as a design equation. If P_{out}^{Av} is the average output power of the oscillator, a minimum value for U^{Av} is prescribed by the relation

$$\nu \cdot U^{Av} > P_{out}^{Av}. \tag{16}$$

Inserting P_{out}^{Av}/ν in (1) for $U_{h\nu}$, together with suitable values for Q and the interaction volume V, gives an indication of the minimum average ac current density which will be required to cause the oscillator to function and deliver the desired output P_{out}^{Av}. The relative importance of the various design parameters and the effects of their variation are also suggested by the way in which they appear in (1).

A deeper significance is revealed if Eq. (1) is modified somewhat so that its frequency character may be studied. If Q_0 is the unloaded quality factor of the resonant structure and if there is to be any useful output whatever, then

$$Q < Q_0, \tag{17}$$

and (1) can be rewritten

$$I_0^2 > (2\omega^2\epsilon/Q_0^2 V)U^{Av}. \tag{18}$$

In scaling down any Group I oscillator to operate at a shorter wavelength, the interaction volume V is proportional to λ_0^3 and Q_0 is proportional to $\lambda_0^{\frac{1}{2}}$. Therefore, (18) becomes

$$I_0^2 > K\nu^6 U^{Av}, \tag{19}$$

in which the constant K depends on the geometry of the resonant structure, the mode of oscillation, the method of excitation, etc.

U^{Av}, the average signal energy stored in the resonant field, has a minimum value governed by the requirement that it be sufficient to override the noise. Thus there is a maximum frequency beyond which the inequality (19) is violated and the oscillator does not function. This maximum frequency depends only on I_0, K, and the temperature (which determines the noise and hence U_{min}^{Av}). Since only $\sqrt[6]{K}$ affects this maximum frequency, altering the shape of the resonant structure has substantially no effect.

For these reasons, (18) may be appropriately labeled the Frequency Limitation Equation. A complete knowledge of how the parameters I_0, Q_0, V, and U^{Av} vary with frequency in a given Group I oscillator would permit an estimation of the frequency bound on that oscillator. Different oscillators may be compared with respect to their suitability for scaling to shorter wavelengths. Thus, (18) may also be used as a yardstick for good design.

MINIMUM U^{Av}

As mentioned previously, there must be sufficient signal energy stored in the field of the resonant structure to insure a dectectable signal in the presence of the noise. The noise arises from two sources: (1) The conductors of the resonant structure are at a finite temperature. Thus the free electrons in these conductors must have a random motion[2,3] which generates a noise field. (2) The beam consists of electrons which were randomly emitted from a heated surface. True, a large velocity has been superimposed on this random motion, but the stochastic velocities are still present in the beam and they also induce a noise field in the resonant structure.

The coupling device used to extract signal energy from the resonant field will detect some of this noise. How much noise the coupling device detects depends on its frequency response. The least amount this can be is the noise which is contained in the single mode that the bunched beam is designed to excite. The minimum average noise energy level is thus taken to be[4,5]

$$U_N^{Av} = (h\nu/e^{h\nu/kT} - 1). \tag{20}$$

Here, the coupler and the resonant structure have been assumed to be at the same equilibrium temperature T and the noise in the beam has not been considered.

U^{Av} should be at least equal to U_N^{Av} to insure a reliable output. Thus

$$U_{min}^{Av} = (h\nu/e^{h\nu/kT} - 1). \tag{21}$$

A further consideration arises. For values of $h\nu \gtrsim 0.692kT$ the expression for U_{min}^{Av} given in (21) is less than one quantum. It is possible to have an average stored energy which is less than this amount by having one quantum of signal energy stored for an appropriate portion of the time and no signal energy stored the rest of the time. However, if coherent oscillations are desired, at least one quantum of signal energy must be stored in the resonant structure continuously. Therefore we take $U_{min}^{Av} = h\nu$ for $h\nu \gtrsim 0.692kT$. These two expressions for U_{min}^{Av} can be stated in terms of the dimensionless quantities

$$u = U_{min}^{Av}/kT, \quad \xi = h\nu/kT$$

as follows:

$$\begin{cases} u = (\xi/e^{\xi} - 1), & 0 \lesssim \xi \lesssim 0.692 \\ u = \xi, & 0.692 \lesssim \xi < \infty. \end{cases} \tag{22}$$

The universal curve of u *versus* ξ is shown in Fig. 1. For a temperature of 300°K, u is less than 1.6 up to a frequency of 10^{13} cps and over this range U_{min}^{Av} as chosen above does not differ greatly from the classical value.

Since $U_{min}^{Av} = ukT$, it would seem at first glance that supercooling the coupler and resonant structure

[2] R. B. Barnes and S. Silverman, Revs. Modern Phys. **6**, 162–192 (1934).
[3] H. Nyquist, Phys. Rev. **32**, 110–113 (1928).
[4] F. K. Richtmyer and E. H. Kennard, *Introduction to Modern Physics* (McGraw-Hill Book Company, Inc., New York, 1947), fourth edition, pp. 171–177.
[5] J. C. Slater, *Introduction to Chemical Physics* (McGraw-Hill Book Company, Inc., New York, 1939), first edition, pp. 142–145 and pp. 313–316.

would allow the minimum required stored energy to be decreased proportionately. However, it must be remembered that no account has been taken of the noise in the beam, and this will not be affected by any supercooling of the structure. Since the beam originates in a cathode which is at a much higher temperature, it may well be that the beam noise is the dominating factor. To the author's knowledge, no method has yet been devised for calculating the effect of the beam noise. However, it does seem safe to say that (22) is a lower bound on the energy stored.

MAXIMUM I_0

A discussion of the factors influencing the attainment of high beam current densities is not a direct concern of this paper. This subject has been given wide treatment elsewhere.[6-9] Suffice it to say that the accelerating voltage and focusing control determine a maximum attainable value for a dc beam. It must be emphasized that the term which occurs in the Frequency Limitation Equation refers to an ac beam current density. It is more difficult to achieve an ac beam of a given density than the corresponding dc beam. This difficulty increases with frequency.

Q_0 AND V

No general expressions can be given for the quality factors and interaction volumes of Group I oscillators, since these quantities depend on the geometric design of a particular oscillator. Two specific types—the klystron and the magnetron—will be discussed. The treatment for other types would proceed along precisely similar lines.

The Klystron

The Two-Gap Klystron

Considering first the two-gap klystron, if the catcher cavity is described in terms of orthogonal coordinates, it is apparent that the beam will only excite TM_{mnp} modes, in which m and n are the transverse mode numbers and p is the axial mode number. For cavities in which the entrant and exit surfaces are parallel and transverse to the beam, setting $p=0$ yields the characteristic operating condition, that is, a uniform electric field across the gap.

For the case of a rectangular cavity of square cross section,[10,11]

$$Q_{mn\cdot 0}=\frac{\eta}{R_s}\cdot\frac{\omega}{c}\cdot\frac{d}{2[1+(2d/a)]}, \tag{23}$$

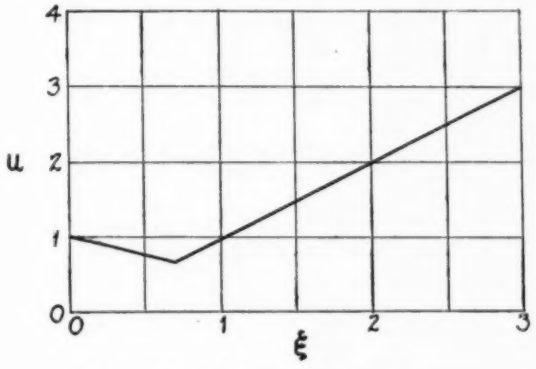


FIG. 1. Universal curve of threshold energy.

in which $a=(\lambda_0/2)(m^2+n^2)^{\frac{1}{2}}$ is the transverse dimension of the cavity, d the axial dimension, λ_0 the free-space wavelength corresponding to the resonant frequency of the $mn0$th mode, η the impedance of free-space, and R_s the surface resistivity of the cavity walls. Regardless of the mode, it is seen from (23) that

$$Q_{mn\cdot 0}<\frac{\eta}{R_s}\cdot\frac{\omega}{c}\cdot\frac{d}{2}. \tag{24}$$

It is further necessary to have $d<(\lambda_0/2)$ for efficient transfer of power from the beam to the field. (This is the well-known transit time effect.) Thus

$$Q_{mn\cdot 0}<(\pi\eta/2R_s) \tag{25}$$

for *all* modes.

The customary operation for klystrons is to have the beam engage only the central space half-cycle of the field, irrespective of the mode. For the beam to react with more than one space half-cycle of the field would require a current phase which reversed from one space half-cycle to the next. This would present a formidable problem in beam bunching. It would result in a larger interaction volume V and hence a higher frequency limit, but seems technologically unfeasible and will not be considered further.

If the beam is confined to the central space half-cycle, the interaction volume V is easy to compute for the rectangular cavity. It is given by

$$V<(\lambda_0^2 d/4)\cdot[(m^2+n^2)/mn], \tag{26}$$

and assuming that the beam excites a symmetric mode $(m=n)$,

$$V<(\lambda_0^3/4), \tag{27}$$

regardless of the value of m. For this case, (18) becomes

$$I_0^2>\frac{2\omega^2\epsilon}{(\pi\eta/2R_s)^2\cdot(\lambda_0^3/4)}U^{\mathrm{Av}}. \tag{28}$$

For copper, $R_s=\frac{1}{4}\cdot 10^{-6}\sqrt{\nu}$, and if a free-space interior is assumed,

$$I_0^2>1.85\cdot 10^{-53}\cdot\nu^6\cdot U^{\mathrm{Av}}. \tag{29}$$

[6] A. V. Haeff, Proc. Inst. Radio Engrs. **27**, 586–602 (1939).
[7] J. R. Pierce, J. Appl. Phys. **10**, 715–724 (1939).
[8] L. P. Smith and P. L. Hartmann, J. Appl. Phys. **11**, 220 (1940).
[9] B. J. Thompson and L. B. Hedrick, Proc. Inst. Radio Engrs. **28**, 318–324 (1940).
[10] S. Ramo and J. R. Whinnery, *Fields and Waves in Modern Radio* (John Wiley & Sons, Inc., New York 1944), first edition, pp. 383–396.
[11] R. I. Sarbacher and W. A. Edson, *Hyper and Ultra-High Frequency Engineering* (John Wiley & Sons, Inc., New York, 1943), first edition, pp. 364–378.

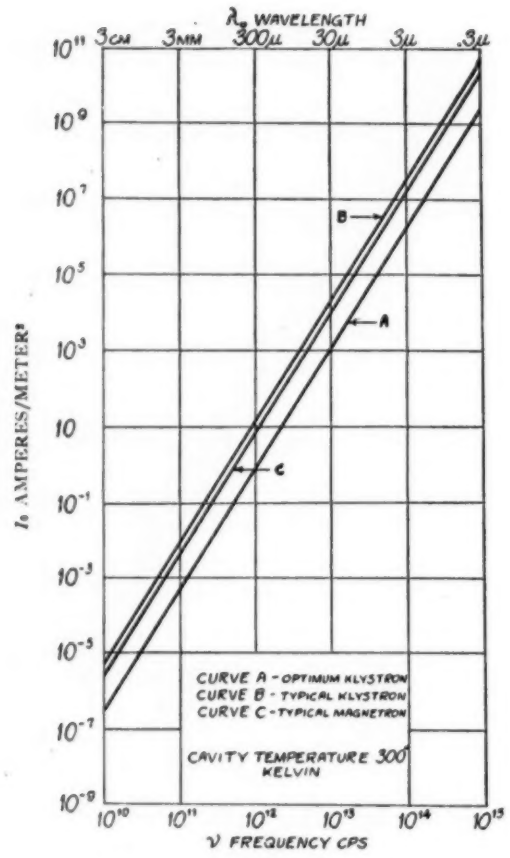


FIG. 2. Minimum average ac beam current density required *vs* desired frequency.

This same procedure may be repeated for a two-gap klystron using cylindrical cavities of circular cross section.[12,13] The results are

$$Q_0 < (\pi\eta/2R_s)$$

for all TM_{m00} modes, and

$$V < 0.23\lambda_0^3$$

for all modes when only the central space half-cycle is engaged. This yields

$$I_0^2 > 2.01 \cdot 10^{-53} \cdot \nu^6 \cdot U^{Av}. \qquad (30)$$

Referring to (19), one sees that $K = 1.85 \times 10^{-53}$ for rectangular cavities; 2.01×10^{-53} for cylindrical cavities. The sixth root of K has not varied greatly, as was anticipated. Small improvement would be expected from using a different cavity design.

If a cavity temperature of 300°K is assumed, U_{min}^{Av} can be computed from Fig. 1 and its value substituted in (30). Thus a minimum average required beam current density will be specified for any frequency. This relationship is shown in curve *A* of Fig. 2.

Scaling of Existing Klystrons

A much more realistic picture is obtained when the values for the interaction volume V of actual tubes are substituted in the Frequency Limitation Equation. From a representative list of commercial klystrons it was determined that $V < 10^{-3}\lambda_0^3$ for all tubes considered. Inserting $V = 10^{-3}\lambda_0^3$ in (18) together with (22) and (25) yields the equation

$$I_0^2 > 0.463 \cdot 10^{-50} \cdot \nu^6 \cdot U_{min}^{Av}. \qquad (31)$$

This defines a bound shown as Curve *B* in Fig. 2. It is to be noted that at any frequency, approximately sixteen times as much current density is required as was found for the optimum calculation.

A Specified Power Output

If, instead of a threshold signal, one requires a sensible output from the klystron, the value of U^{Av} must be raised correspondingly. Since all of the stored energy is not to be extracted every cycle,

$$\nu \cdot U^{Av} > P_{out}^{Av}, \qquad (32)$$

and for this case, (31) transforms to be

$$I_0^2 > 0.463 \cdot 10^{-50} \cdot \nu^5 P_{out}^{Av}. \qquad (33)$$

The bound defined by (33) is plotted in Fig. 3 for several values of output.

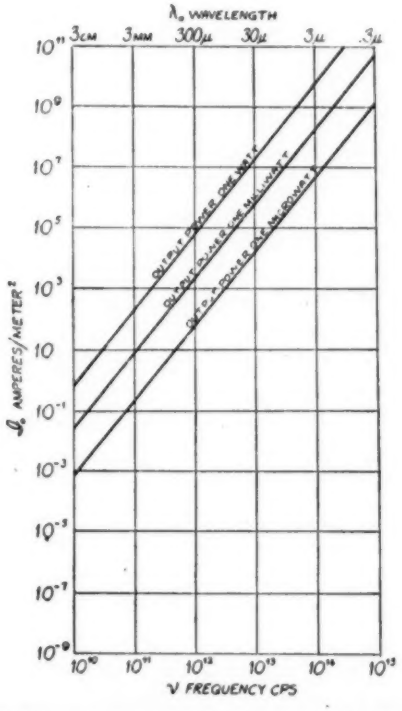


FIG. 3. Minimum average ac beam current density required *vs* frequency and desired output for typical klystron.

[12] S. Ramo and J. R. Whinnery, *Fields and Waves in Modern Radio* (John Wiley & Sons, Inc., New York, 1944), first edition, pp. 396–404.

[13] R. I. Sarbacher and W. A. Edson, *Hyper and Ultra-High Frequency Engineering* (John Wiley & Sons, Inc., New York, 1943), first edition, pp. 378–387.

Extension to Reflex Klystrons and Multiple Cavity Klystrons

In the preceding analysis, the buncher cavity was involved only in that it was assumed that the beam was bunched. The limitations arose entirely due to the action at the catcher. Thus the whole analysis applies equally well to reflex klystrons. Due to the added collision problem, when high beam current densities are being considered, the limitations on reflex klystrons will be more severe than the limitations on two-gap klystrons.

It is well known that the beam, on emerging from the catcher of a two-gap klystron is still partially bunched. If a second catcher is added to extract more energy from the beam, the over-all efficiency is improved. This idea can be extended to include n catchers in tandem. However, as far as the frequency bound is concerned, it will be of no avail to use additional catchers. It has already been shown that if the first catcher is oscillating at its maximum frequency, the power supplied to it by the beam is just enough to overcome the losses. The partially de-bunched beam will not be this successful in the second catcher and therefore its signal-to-noise ratio will be poorer and it will contribute no useful output. This same remark can be made more and more strongly for successive catchers. The frequency bound for a two-gap klystron is therefore also the frequency bound for a multi-gap klystron.

The Magnetron

There are so many interrelated parameters governing the proper geometry of a magnetron that no statement can be made as to the values these parameters should assume for optimum design. However, the vast wealth of experimental information leading to the good designs now commercially available may be utilized for scaling. If the interaction volume is taken to be the total volume between anode and cathode divided by the number of resonators, the calculation of V is a simple matter. For all of the commercial magnetrons for which data was available it was found that $V<10^{-2}\lambda_0^3$.

Establishing a value for Q_0 of one of the resonators is a difficult task because the shapes of the resonators vary so widely from one magnetron to the next, and because further there is usually some coupling between resonators. However, a good approximation can be obtained by using as a model a cylindrical cavity oscillating in its TE_{111} mode. Moreno[14] gives for the unloaded Q for such a cavity

$$Q_0=\frac{z_0}{\delta}\cdot\frac{[(1.841)^2+(\pi a/2z_0)^2]\cdot[1-(1/(1.841)^2)]}{[(1.841)^2\cdot z_0/a+(a^2\pi^2/4z_0^2)+(a(z_0-a)/4z_0^2)\cdot(\pi^2/(1.841)^2)]}, \quad (34)$$

[14] T. Moreno, *Microwave Transmission Design Data* (McGraw-Hill Book Company, Inc., New York, 1948), first edition, p. 219.

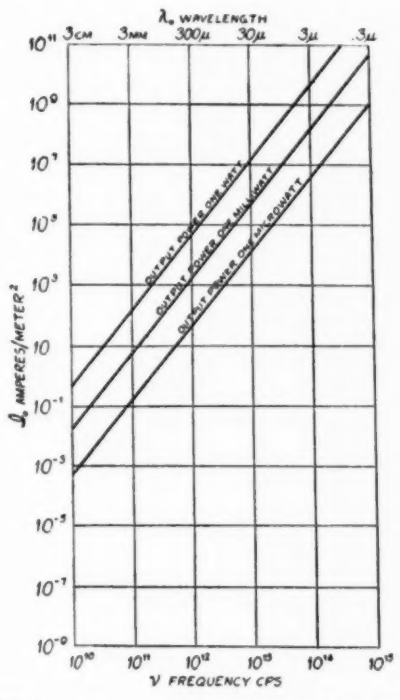


FIG. 4. Minimum average ac beam current density required *vs* frequency and desired output for typical magnetron.

in which

a = radius of the cylinder,
$2z_0$ = height of the cylinder,
δ = skin depth of the conductor.

For a given frequency, this expression for Q_0 is a slowly varying function of $2z_0/a$. The central value is $\pi/4\cdot\eta/R_s$ and we shall assume that this approximates all magnetrons well enough for the purpose of calculations.

If $V=10^{-2}\lambda_0^3$ and $Q_0=\pi/4\cdot\eta/R_s$ are substituted in the Frequency Limitation Equation, there results

$$I_0^2>1.85\cdot10^{-51}\cdot\nu^5\cdot U_{\min}{}^{Av}. \quad (35)$$

The bound defined by (35) is shown as Curve C in Fig. 2. The bounds for typical klystrons and magnetrons are not too much different from each other.

As in the case of the klystron, it is possible to express U^{Av} in terms of a desired power output by the relation (32). For this case, (35) is replaced by

$$I_0^2>1.85\cdot10^{-51}\cdot\nu^5\cdot P_{out}{}^{Av}. \quad (36)$$

The frequency bound determined by (36) is plotted in Fig. 4 for several values of output.

CONCLUSIONS

It has been shown that all Group I oscillators are limited by the requirement that the inequality

$$I_0^2>\frac{2\omega^2\epsilon}{Q_0^2V}\cdot U^{Av} \quad (18)$$

must be satisfied. For all practical purposes, there is not much that can be done in the way of design to improve Q_0 and V over the values presently attainable, so one may argue that

$$I_0^2 > K\nu^6 U^{Av}, \tag{19}$$

where K has been optimized about as well as can be expected.

For a given temperature U^{Av} has a minimum value established by the noise level. Thus, essentially, the frequency bound is set by the current density of the beam. Only by achieving higher beam densities would it seem reasonable to expect any significant increase in the frequencies obtainable.

Referring to Fig. 2, we observe that with an ac beam current density of 1 ampere/cm^2, one would not expect to generate in a magnetron a wavelength shorter than about 30 microns. The assumptions made in deriving this bound were generous enough so that it is perhaps off by one magnitude or more. Should this prove to be the case, it would suggest that experimenters have just about reached the frequency limit with present techniques. It would suggest further that there are only two favorable lines of attack. One would be to concentrate on developing denser beams. The other would be to abandon Group I oscillators and investigate some of the more recent schemes for producing coherent oscillations which do not require a resonant energy extractor, such as the traveling wave tube‡ and devices in which a bunched beam radiates directly.[15,16]

ACKNOWLEDGMENT

The author is indebted to Dr. Lloyd T. De Vore who first suggested this problem, and to Professors W. L. Everitt, H. von Foerster, E. C. Jordan, C. Nash, and N. Wax for their generous counsel.

‡ There is some reason to believe that the foregoing analysis can be extended to traveling wave tubes by changing to a moving coordinate system. In a sense, the traveling wave tube has a distributed resonant energy-extractor for which a Q_0 and V can be computed. Due to its distributed characteristics, it offers more promise as a coherent generator of short wavelengths than Group I oscillators. The author hopes to consider this problem in a later paper.

[15] P. D. Coleman, "Generation of Millimeter Waves," Ph.D. thesis, Massachusetts Institute of Technology, May, 1951.

[16] H. Motz, J. Appl. Phys. **22**, 527 (1951).

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 8 AUGUST, 1952

A Mathematical Analysis of a Dielectric Amplifier

LOUIS A. PIPES
Department of Engineering, University of California, Los Angeles, California

(Received April 1, 1952)

This paper presents a mathematical analysis of the fundamental circuit on which the operation of dielectric amplifiers depends. The analysis is based on the assumption that the effective hysteresis curve of the amplifier's dielectric material may be represented by a hyperbolic sine function. The case of resistive load is analyzed and expressions for the steady-state input and output currents are calculated. From a consideration of the transient response of the amplifier, an estimate of its time constant is obtained.

I. INTRODUCTION

A GREAT deal of the important progress made in all phases of electrical engineering for the past thirty years is due to the successful use of circuitry based on the vacuum tube. Despite the excellent performance of the equipment based on this type of circuitry, its use is limited because the tube is a very delicate device. The tube itself has a relatively short life and is sensitive to overloads, shocks, and vibrations. This limits the reliability of this type of electronic equipment to that of the vacuum tube and makes its use, particularly in military operations, hazardous. This lack of complete reliability has encouraged the study of the theory and possible application of tubeless devices. It appears that the most promising competitors of the vacuum tube today are the transistor, the resistance or crystal amplifier, and the magnetic and dielectric amplifier. The transistor and the magnetic amplifier have been covered quite extensively in recent literature.[1,2]

The dielectric amplifier is a relatively new device from the standpoint of applications.[3] It has the advantages of ruggedness, compactness, efficiency, reliability, and high gain. It is less expensive than tube or magnetic amplifiers and is particularly adaptable to high impedance control circuits. It has the disadvantages of gain drift, and is often noisy as a result of molecular disturbances sometimes found in its dielectric. Dielectric amplifiers seem particularly well-suited for use in regulators, relays, limiters, servo systems, phase shifters, modulators, and multivibrators.

[1] R. L. Wallace, Bell System Tech. J. **30**, 381 (1951).

[2] J. G. Miles, "Bibliography of Magnetic Amplifier Devices and the Saturable Reactor Art," Am. Inst. Elec. Engrs. Technical Paper 51-388. (This is a comprehensive bibliography concerning Magnetic Amplifiers and their practical application to the middle of the year 1951.)

[3] A. M. Vincent, Electronics **24**, Part II.84 (1951).

It is difficult to credit any one individual with the invention of the dielectric amplifier. However, the discovery of the phenomenon that certain dielectrics have nonlinear characteristics naturally suggested their use in amplifiers. As early as 1912, Debye[4] discovered that when atoms having different electron affinities combined, electric dipoles were formed. The nonlinearity of the magnetization curves of materials containing iron is usually termed a "ferro-magnetic effect." This effect is known to depend on the parallel orientation of electron spins of the magnetic dipoles present in the magnetic material. By analogy to the magnetic case, nonlinearity of the dielectric constant, caused by the alignment of electric dipoles of atomic ions in certain nonmagnetic crystalline structures, has been termed the "ferro-electric effect," although there is no iron involved.

Many substances that exhibit the ferro-electric effect have been discovered. Rochelle salt was found to be a nonlinear dielectric by Valasek[5] in 1921. The ferro-electric properties of the dihydrogen phosphates and aresenates were studied in 1935 by Busch and Scherrer.[6] Wainer and Salomon[7] discovered the unusual dielectric properties of barium titanate in 1942. The ferro-electric properties of the titanates and some of their applications have been studied by A. von Hippel and his associates at the Massachusetts Institute of Technology.[8,9]

The barium titanate compositions that are frequently used today in dielectric amplifiers are synthetic crystalline materials which have a maximum dielectric constant in the range 1500 to 10,000. The dielectric constant of barium titanate changes greatly with temperature variations. The temperature at which the dielectric constant is a maximum is called the Curie point. It is possible to shift the location of the Curie point of barium titanate by the addition of strontium titanate. The greatest gain of dielectric amplifiers is obtained when they are operated at the Curie point of their dielectric material. Usually they are designed to be operated at a temperature just above the Curie point. It has been found possible to compensate for the change in gain due to temperature variations by using temperature control devices, or a combination of two dielectrics, one operating above its Curie point and the other one below it. The dielectric constants of some materials such as barium-lead-zirconate remain approximately constant for a considerable temperature variation. When these materials are used, temperature control is not a serious consideration.

[4] P. Debye, *Polare Molekeln* (S. Hirzel, Leipzig, Germany, 1929).
[5] J. Valasek, Phy. Rev. **17**, 475 (1921).
[6] G. Busch, Phys. Acta. **11**, 269 (1938).
[7] E. Wainer, Trans. Am. Electrochem. Soc. **89**, 138 (1946).
[8] A. von Hippel, ONR Research Reviews **24**, No. 3 (1951).
[9] A. von Hippel, *Dielectric Relaxation Phenomena in Liquids and Solids* (Division of Solid State Physics; Am. Phys. Soc. Pittsburgh, March 8–10, 1951).

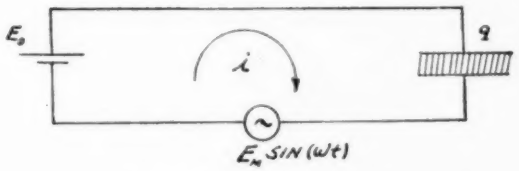


FIG. 1. Fundamental circuit.

II. THE BASIC PRINCIPLES OF DIELECTRIC AMPLIFIER OPERATION

The fundamental principles underlying the operation of a dielectric amplifier may be illustrated by means of Fig. 1. The elementary dielectric amplifier of Fig. 1 can be compared to a circuit in which the variations of the grid potential of a vacuum tube control a relatively large amount of power in the plate circuit. This elementary dielectric amplifier consists of a capacitor in series with a source of harmonic electromotive force and a constant potential source. The amplitude of the current that flows in the circuit is controlled by varying the impedance of the capacitor. The variation of the impedance of the capacitor is provided by changing the degree of saturation of its dielectric by varying the magnitude of the direct potential E_0.

The dielectric constant of a nonlinear dielectric material *decreases* with its degree of saturation. The capacitance, and, therefore, the *admittance* of a nonlinear capacitor is diminished with saturation. Therefore the current in the circuit may be reduced by saturating the dielectric of the capacitor.

The Saturation Curve of the Nonlinear Capacitor

In order to study the behavior of the circuit of Fig. 1 mathematically, it is necessary to represent the saturation curve $V(q)$, expressing the potential drop across the plates of the capacitor when it carries a charge q, by an analytical expression. A typical saturation curve of a nonlinear capacitor has the general form given by Fig. 2.

The typical saturation curve $V(Q)$ may be represented by a variety of analytical expressions. A very useful one that can be adjusted to fit the usual experi-

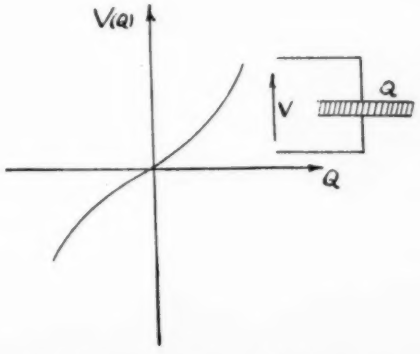


FIG. 2. Saturation curve of the capacitor.

mental $V(Q)$ curve closely is the hyperbolic sine curve

$$V(q)=\frac{S_0}{a}\sinh(aq). \tag{2.1}$$

In this expression, S_0 is the *initial* elastance of the capacitor defined by

$$S_0=\left(\frac{dV}{dq}\right)_{q=0}. \tag{2.2}$$

This is the reciprocal of the *initial* capacitance defined by

$$C_0=\left(\frac{dq}{dV}\right)_{q=0}. \tag{2.3}$$

The constant a is a measure of the nonlinearity of the saturation curve of the capacitor. The slope of the $V(q)$ curve is given by

$$\frac{dV}{dq}=S_0\cosh(aq)=V'. \tag{2.4}$$

Therefore the curve $V(q)$ has constant slope and is

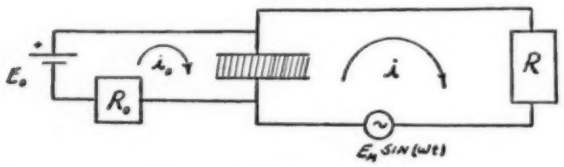


FIG. 3. Dielectric amplifier.

therefore a straight line only when $a=0$. If (2.4) is divided by (2.1), the result may be written in the form

$$a=\frac{V'}{V}\tanh(aq). \tag{2.5}$$

For large values of q, $\tanh(aq)=1$. Therefore (2.5) may be written in the form

$$a=\left(\frac{V'}{V}\right)_{q\to\infty}. \tag{2.6}$$

Equation (2.6) may be used to obtain an estimate for a from the empirical saturation curve of the nonlinear capacitor.

In order to accentuate the most important features of the elementary dielectric amplifier of Fig. 1 and to keep the mathematical complexity of the analysis to a minimum, let the resistance of the circuit be neglected.

The equation of the circuit may be obtained by equating the applied potential of the circuit to the potential drop across the plates of the capacitor $V(q)$ in the form

$$V(q)=\frac{S_0}{a}\sinh(aq)=E_0+E_m\sin(wt)=E(t). \tag{2.7}$$

Therefore,

$$q=\frac{1}{a}\sinh^{-1}\left(\frac{aE}{S_0}\right). \tag{2.8}$$

If (2.8) is differentiated with respect to t, the following expression for the current i is obtained:

$$i(t)=(a^2E^2+S_0^2)^{-\frac{1}{2}}dE/dt. \tag{2.9}$$

If it is assumed that the control potential, E_0, is much greater than the maximum value of the applied harmonic potential of the circuit, so that

$$E_0\gg E_m, \tag{2.10}$$

then the expression (2.9) may be written in the form,

$$i(t)=A_0\cos(wt), \tag{2.11}$$

where

$$A_0=wE_m(a^2E_0^2+S_0^2)^{-\frac{1}{2}}. \tag{2.12}$$

A measure of the gain of the dielectric amplifier is

$$\mu_0=\left(\frac{\partial A_0}{\partial E_0}\right)_{E_0}=-wa^2E_mE_0(a^2E_0^2+S_0^2)^{-\frac{1}{2}}. \tag{2.13}$$

A small increase ΔE_0 of the control potential is accompanied by a change of ΔA_0 in the amplitude of the current given by

$$\Delta A_0=\mu_0\Delta E_0. \tag{2.14}$$

Since the gain of the amplifier μ_0 is negative, a slight *increase* of the control potential corresponds to a *decrease* in the amplitude of the current in the amplifier.

III. A PRACTICAL DIELECTRIC AMPLIFIER CIRCUIT NOTATION

Having outlined the basic principles involved in the operation of dielectric amplifiers, a basic circuit of fundamental importance in the design of dielectric amplifiers will now be studied. This circuit is illustrated by Fig. 3.

The more elaborate dielectric amplifier circuits suggested for practical applications are essentially variations and combinations of this basic circuit. Because of its technical importance a mathematical analysis of its operation will now be presented. The following notation will be used:

$q_0(t)$ = the circulating charge in the input circuit.
$q(t)$ = the circulating charge in the output circuit.
$i_0(t)=\dot{q}_0$ = the input current.
$i(t)=\dot{q}$ = the output current.
S_0 = the *initial* charge of the nonlinear capacitor.
C_0 = the *initial* capacitance of the nonlinear capacitor.
R_0 = the resistance of the input circuit.
R = the resistance of the output circuit.
E_0 = the control or input potential.

$E_m \sin(wt)$ = the potential applied to the output circuit.
$Q(t) = (q_0 - q)$ = the charge of the capacitor.
$V(Q)$ = the potential between the plates of the nonlinear capacitor.
$V(Q) = (S_0/a)\sinh(aQ)$ = an analytical expression for the saturation curve of the nonlinear capacitor.
a = a constant determined from the empirical saturation curve of the nonlinear capacitor.
$R_e = R_0R/(R_0+R)$ = the effective resistance of the circuit.
$b = 1/R_eC_0$ = a constant.
$T_0 = 1/b = R_eC_0$ = the initial time constant of the circuit.
$Q_0(t)$ = the first approximation to the charge of the capacitor.
$A = E_0C_0R_ea/R_0$ = a constant.
$B = E_ma/R((b^2+w^2))^{\frac{1}{2}}$ = a constant.
$I_n(x)$ = the modified Bessel function of the first kind of order n and argument x.

The General Equations of the Dielectric Amplifier

If Kirchoff's electromotive force law is applied to the two loops of the circuit of Fig. 3, the following equations are obtained:

$$R_0i_0 + V(Q) = E_0, \tag{3.1}$$

$$Ri - V(Q) = E_m \sin(wt), \tag{3.2}$$

where

$$Q = (q_0 - q) = \int_0^t (i_0 - i)\,dt. \tag{3.3}$$

$V(Q)$ is the potential drop across the plates of the nonlinear capacitor, and q_0 and q are the circulating charges in the two loops of the dielectric amplifier circuit.

Equations (3.1) and (3.2) may be written in the following form:

$$i_0 + V(Q)/R_0 = E_0/R_0 \tag{3.4}$$

and

$$i - V(Q)/R = E_m \sin(wt)/R. \tag{3.5}$$

If (3.5) is subtracted from (3.4), the result is

$$(i_0 - i) + \left(\frac{1}{R_0} + \frac{1}{R}\right)V(Q) = E_0/R_0 - E_m \sin(wt)/R \tag{3.6}$$

or

$$\dot{Q} + \frac{V(Q)}{R_e} = \frac{E_0}{R_0} - \frac{E_m}{R}\sin(wt), \tag{3.7}$$

where

$$R_e = R_0R/(R_0+R). \tag{3.8}$$

Equation (3.7) is a nonlinear differential equation for the determination of $Q(t)$. When Q has been found by solving (3.7), then the input current i_0 may be determined by means of (3.1) in the form

$$i_0(t) = [E_0 - V(Q)]/R_0. \tag{3.9}$$

Similarly, as a consequence of (3.2), the output current i is given by

$$i(t) = [E_m \sin(wt) + V(Q)]/R.$$

In order to effect the solution of (3.7) it is necessary to represent the saturation curve $V(Q)$ by an analytical expression. In this analysis, the expression given by (2.1) will be assumed. With this assumption, Eq. (3.7) takes the form

$$DQ + \frac{S_0}{R_ea}\sinh(aQ) = \frac{E_0}{R_0} - \frac{E_m}{R}\sin(wt), \quad D = \frac{d}{dt}.$$

If the term Q_0/C_0R_e is added to both members of (3.11), the result is

$$(D+b)Q(t) = \frac{E_0}{R_0} - \frac{E_m}{R}\sin(wt) + \frac{b}{a}[aQ - \sinh(aQ)].$$

If it is assumed that the circuit is inert at $t=0$, then $Q(0)=0$ and the differential Eq. (3.12) may be written as a nonlinear integral equation having the following form:[10,11]

$$Q(t) = Q_0(t) + \frac{b}{a}\int_0^t e^{-b(t-u)}\{aQ(u) - \sinh[aQ(u)]\}\,du,$$

where $Q_0(t)$ is the solution of the linear differential equation

$$(D+b)Q_0(t) = E_0/R_0 - E_m \sin(wt)/R,$$

with $Q_0(0)=0$.

The nonlinear integral Eq. (3.13) is a special case of an integral equation studied by Lalesco.[12] Its solution may be effected by setting up the following sequence of functions:

$$\begin{aligned}
Q_1(t) &= Q_0(t) + \frac{b}{a}\int_0^t e^{-b(t-u)} \\
&\qquad \times \{aQ_0(u) - \sinh[aQ_0(u)]\}\,du, \\
Q_2(t) &= Q_0(t) + \frac{b}{a}\int_0^t e^{-b(t-u)} \\
&\qquad \times \{aQ_1(u) - \sinh[aQ_1(u)]\}\,du, \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
Q_{k+1}(t) &= Q_0(t) + \frac{b}{a}\int_0^t e^{-b(t-u)} \\
&\qquad \times \{aQ_k(u) - \sinh[aQ_k(u)]\}\,du.
\end{aligned} \tag{3.15}$$

[10] L. A. Pipes, "Operational Methods in Nonlinear Mechanics," Report 51-10 (Department of Engineering, University of California, Los Angeles, 1951).

[11] L. A. Pipes, *Applied Mathematics for Engineers and Physicists* (McGraw-Hill Book Company, Inc., New York, 1946), Chapter XXI.

[12] Vito Volterra, *Leçons sur les Equations Integrales* (Gauthier-Villars, Paris, France, 1913), p. 90.

The limit of the sequence of functions, $Q_0(t)$, $Q_1(t)$, $\cdots$, $Q_k(t)$, defined by (3.15), is the solution of (3.13). That is,

$$\lim_{k\to\infty} Q_k(t)=Q(t). \tag{3.16}$$

The function $Q_0(t)$ is the solution of (3.14) with $Q_0(0)=0$; it is given by

$$Q_0(t)=E_0\frac{(1-e^{-bt})}{R_0b}-\frac{E_m}{R}w\left[\frac{e^{-bt}}{(b^2+w^2)}+\frac{\sin(wt-\phi)}{w(b^2+w^2)}\right] \tag{3.17}$$

$$\tan\phi=w/b.$$

This is the variation of the charge on the plates of the capacitor if it had a constant capacitance C_0. It is the first approximation to the solution of the integral Eq. (3.13). An investigation of the convergence of the sequence (3.15) indicates that it converges rapidly, particularly for small values of a.

IV. THE THEORY OF THE FIRST APPROXIMATION

The response of the dielectric amplifier based upon the first approximation to the solution of the integral Eq. (3.13) may be obtained by substituting $Q_0(t)$ into the Eqs. (3.9) and (3.10) for the input and output currents of the amplifier and thus obtaining

$$i_0(t)=1/R_0\{E_0-S_0\sinh[aQ_0(t)]/a\} \tag{4.1}$$

and

$$i(t)=\frac{1}{R}\left\{E_m\sin(wt)+\frac{S_0}{a}\sinh[aQ_0(t)]\right\}. \tag{4.2}$$

A. The Effective Time Constant of the Amplifier

It can be seen from the form of the transient solution for $Q_0(t)$ given by (4.12) that the quantity

$$T_0-\frac{1}{b}=R_eC_0=\frac{R_e}{S_0} \tag{4.3}$$

is a measure of the rapidity with which transient disturbances in the dielectric amplifier circuit attenuate. S_0 is the *initial* elastance of the nonlinear capacitor defined by

$$S_0=(dV/dQ)_{Q=0}, \tag{4.4}$$

where

$$V(Q)=S_0\sinh(aQ)/a \tag{4.5}$$

is the saturation curve of the nonlinear capacitor. A better estimate of the response time of the dielectric amplifier may be obtained if the *incremental* elastance of the nonlinear capacitor is used instead of S_0. The incremental elastance of the nonlinear capacitor is defined by the equation,

$$S_i=(dV/dQ)_{Q=Q}=S_0\cosh(aQ). \tag{4.6}$$

Then if Q_e is the mean operating value of Q when deviations from the steady-state operation are initiated, an effective value of the time constant of the amplifier T_e may be defined by the expression

$$T_e=\frac{R_e}{(dV/dQ)_{Q=Q_e}}=\frac{R_e}{S_0\cosh(aQ_e)}. \tag{4.7}$$

For any case other than $Q_e=0$, we have $T_e<T_0$, that T_0 is a measure of the maximum time of respon of the circuit.

B. The Steady-State Response

The steady-state response of the amplifier based on the first approximation may be obtained by substituting the steady-state part of the solution for $Q_0(t)$ into Eqs. (4.17) and (4.18). The steady-state value of $Q_0(t)$, $Q_{0s}(t)$, is given by (4.12) after the transient terms have decayed to zero. It is

$$Q_{0s}(t)=C_0E_0\frac{R_e}{R_0}-E_m\frac{\sin(wt-\phi)}{R(b^2+w^2)^{\frac{1}{2}}},\quad \tan\phi=wR_eC_0. \tag{4.8}$$

In this expression, the quantity Q_e given by

$$Q_e=E_0C_0\frac{R_e}{R_0} \tag{4.9}$$

may be taken to be the mean operating value of Q to be used in an estimate of the effective time constant T_e given by (4.7).

In the usual application the condition

$$wR_eC_0\gg 1 \tag{4.10}$$

is satisfied, and therefore, $\phi\approx\pi/2$. Hence (4.8) may be written in the form

$$Q_{0s}(t)=\frac{[A+B\cos(wt)]}{a}, \tag{4.11}$$

where

$$A=E_0C_0\frac{R_e}{R_0}a \quad\text{and}\quad B=\frac{aE_m}{R(b^2+w^2)^{\frac{1}{2}}}. \tag{4.12}$$

The steady-state input and output currents are given by Eqs. (4.1) and (4.2) in the form

$$i_{0s}(t)=\frac{1}{R_0}\left\{E_0-\frac{S_0}{a}\sinh[A+B\cos(wt)]\right\} \tag{4.13}$$

and

$$i_s(t)=\frac{1}{R}\left\{E_m\sin(wt)+\frac{S_0}{a}\sinh[A+B\cos(wt)]\right\}. \tag{4.14}$$

The harmonic content of the input and output currents will now be determined. By the addition theorem of the hyperbolic sine we have

$$\sinh[A+B\cos(wt)]=\sinh(A)\cosh[B\cos(wt)]+\cosh(A)\sinh[B\cos(wt)]. \tag{4.15}$$

The following identities are established in works on Bessel functions:[13]

$$\cosh[B\cos(wt)]=I_0(B)+2\sum_{n=1}^{\infty}I_{2n}(B)\cos(2nwt) \quad (4.16)$$

and

$$\sinh[B\cos(wt)]=2\sum_{n=0}^{\infty}I_{(2n+1)}(B)\cos[(2n+1)wt], \quad (4.17)$$

where $I_n(x)$ is the modified Bessel function of the first kind of order n and argument x. These functions have been extensively tabulated.[14] By the use of the identities (4.15), (4.16), and (4.17), the expressions (4.13) and (4.14) for the steady-state input and output currents may be written in the following form:

$$i_{0s}(t)=\frac{E_0}{R_0}-\frac{S_0}{aR_0}\sinh(A)\left[I_0(B)+2\sum_{n=1}^{\infty}I_{(2n)}(B)\cos(2nwt)\right]-\frac{2S_0}{aR_0}\cosh(A)\times\left[\sum_{n=0}^{\infty}I_{(2n+1)}(B)\cos(2n+1)wt\right] \quad (4.18)$$

and

$$i_s(t)=\frac{E_m}{R}\sin(wt)+\frac{S_0}{aR}\sinh(A)\left[I_0(B)+2\sum_{n=1}^{\infty}I_{(2n)}(B)\cos(2nwt)\right]+\frac{2S_0}{aR}\cosh(A)\times\left[\sum_{n=0}^{\infty}I_{(2n+1)}(B)\cos(2n+1)wt\right]. \quad (4.19)$$

It is thus seen that the steady-state output and input currents contain even and odd harmonics of the fundamental applied potential. If greater accuracy than that given by the first approximation is desired, the second term, $Q_1(t)$, of the sequence (3.15) may be calculated and substituted into the Eqs. (3.9) and (3.10).

V. AN ALTERNATIVE METHOD

If the control potential is much greater than the amplitude of the harmonic potential so that the condition

$$E_0\gg E_m \quad (5.1)$$

is satisfied, then the following alternative method may be used to determine the amplitude of the output current.

Equation (3.7) may be written in the form

$$\frac{S_0}{aR_e}\sinh(aQ)=\frac{E_0}{R_0}-\frac{E_m}{R}\sin(wt)-\dot{Q}. \quad (5.2)$$

An approximate steady-state solution of this equation may be obtained by taking $\dot{Q}$ to be given by $\dot{Q}_{0s}$ and writing it in the following form:

$$\frac{S_0}{aR_e}\sinh(aQ)=\frac{E_0}{R_0}-\frac{E_m}{R}\sin(wt)-\dot{Q}_{0s}. \quad (5.3)$$

The value of Q_{0s} is given by (4.11); it is

$$Q_{0s}=\frac{E_0}{S_0}\frac{R_e}{R_0}+\frac{E_m}{R}\frac{\cos(wt)}{(b^2+w^2)^{\frac{1}{2}}}. \quad (5.4)$$

If this is differentiated and substituted into (5.3), the following equation is obtained:

$$\frac{S_0}{aR_e}\sinh(aQ)=\frac{E_0}{R_0}+P\sin(wt)=I(t); \quad (5.5)$$

where

$$P=\frac{E_m}{R}\left[\frac{w}{(b^2+w^2)^{\frac{1}{2}}}-1\right]. \quad (5.6)$$

If (5.5) is solved for Q the result is

$$Q=\frac{1}{a}\sinh^{-1}\left[\frac{aI}{b}\right]. \quad (5.7)$$

By differentiating (5.7) with respect to t one obtains

$$\dot{Q}=(a^2I^2+b^2)^{-\frac{1}{2}}dI/dt. \quad (5.8)$$

Since

$$I(t)=(E_0/R_0)+P\sin(wt), \quad (5.9)$$

$$dI/dt=wP\cos(wt), \quad (5.10)$$

and hence substituting (5.10) into (5.8) we obtain

$$\dot{Q}=\frac{wP\cos(wt)}{(a^2I^2+b^2)^{\frac{1}{2}}}. \quad (5.11)$$

Now if (3.1) and (3.2) are added together the following result is obtained:

$$i_0R_0+iR=E_0+E_m\sin(wt). \quad (5.12)$$

From the definition of Q given by (3.3), we have

$$i_0-i=\dot{Q}. \quad (5.13)$$

If Eqs. (5.12) and (3.13) are solved simultaneously for i_0 and i, the results are

$$i_0(t)=\frac{E_0+E_m\sin(wt)+R\dot{Q}}{(R+R_0)} \quad (5.14)$$

[13] N. W. McLachlan, *Bessel Functions for Engineers* (Oxford University Press, London, England, 1934).

[14] E. Jahnke and F. Emde, *Tables of Functions* (Dover Publications, New York, 1943), pp. 232–233.

and

$$i(t)=\frac{E_0+E_m\sin(wt)-R_0\dot{Q}}{(R+R_0)}. \tag{5.15}$$

Hence by eliminating $\dot{Q}$ by means of (5.11), the following expression for the steady-state output current is obtained:

$$i_s(t)=\frac{E_0}{(R+R_0)}+\frac{1}{(R+R_0)}\times\left[E_m\sin(wt)-R_0\frac{wP\cos(wt)}{(a^2I^2+b^2)^{\frac{1}{2}}}\right]. \tag{5.16}$$

Now if $E_0\gg E_m$,

$$I\approx E_0/R_0, \tag{5.17}$$

since the term $P\sin(wt)$ in (5.9) can then be neglected. Equation (5.16) can be written in the form

$$i_s(t)=\frac{E_0}{(R+R_0)}+A_0\sin(wt-\phi), \tag{5.18}$$

where

$$A_0=\frac{(E_m^2+K_0^2)^{\frac{1}{2}}}{(R+R_0)},\quad \tan\phi=\frac{K_0}{E_m} \tag{5.19}$$

and

$$K_0=R_0wP\left(\frac{a^2E_0^2}{R_0^2}+b^2\right)^{-\frac{1}{2}}. \tag{5.20}$$

As can be seen from (5.19) and (5.20), the effect of increasing the control potential E_0 is to decrease K_0 and hence to decrease the amplitude of the harmonic load current A_0. The effective gain of the amplifier may be defined by the expression

$$\mu=\partial A_0/\partial E_0. \tag{5.21}$$

If the relative harmonic content of the output current is desired, it may be obtained by substituting the complete expression for $I(t)$ as given by (5.9) into (5.16) and obtaining

$$i_s(t)=\frac{E_0}{(R+R_0)}+\frac{1}{(R+R_0)}\left\{E_m\sin(wt)-\frac{R_0wP\cos(wt)}{[a^2(E/R_0+P\sin wt)^2+b^2]^{\frac{1}{2}}}\right\}. \tag{5.22}$$

The binomial expansion of the radical term in (5.22) yields the required harmonic terms.

The method of analysis discussed in this section is a generalization of the one presented in Sec. II generalized to take into account the effects of resistance. It gives good quantitative results particularly when the control potential E_0 has a value much larger than the maximum harmonic potential E_m.

VI. CONCLUSION

A mathematical method for determining the response of a fundamental dielectric amplifier circuit has been developed. Although the analysis is based on the assumption that the saturation curve of the nonlinear capacitor of the amplifier has the form $V(q)=(S_0/a)\sinh(aq)$, it can easily be extended to cases where other analytical expressions are given for the saturation curve $V(q)$. The solution of the fundamental nonlinear differential Eq. (3.7) may be effected either by an iteration process similar to that of (3.15) or by a method such as that discussed in Sec. V for whatever analytical expression may be used to represent the saturation curve of the capacitor. Once this solution is obtained the amplification constant, time constant, and steady-state response of the amplifier may be readily computed.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 8 AUGUST, 1952

The Emission of Radiation from Diatomic Gases. IV. Emissivity Calculations for CO and HCl for Nonoverlapping Rotational Lines as a Function of Temperature and Optical Density*

S. S. PENNER
Guggenheim Jet Propulsion Center, California Institute of Technology, Pasadena, California
(Received February 4, 1952)

Simplified expressions have been developed for the engineering emissivity of uniformly distributed diatomic gases for nonoverlapping rotational lines with a resonance contour. Unfortunately the rotational half-widths for spectral lines arising from transitions between excited vibrational energy levels are generally not known. For this reason it was necessary to make the assumption that the rotational half-widths for transitions of the form $n \rightarrow n+v$, $v=1$ or 2, are identical. The theoretical analysis is, however, sufficiently general to be useful without modification when accurate data concerning the dependence of line-width on vibrational excitation become available. Explicit expressions have been obtained for the contributions to the total emissivity associated with individual vibration-rotation bands.

Representative emissivity calculations have been carried out for CO and HCl. Comparison of the calculated emissivities of CO with experimental data shows only fair agreement, suggesting either that the assumed description of rotational half-widths is inadequate or else that the empirical emissivity data are not reliable at elevated temperatures.

I. INTRODUCTION

IN previous publications we have described methods for the theoretical calculation of engineering emissivities of diatomic gases from spectroscopic data. Useful approximate methods have been obtained for the limiting cases of (a) extensive overlapping between rotational lines[1,2] and (b) for separated rotational lines without overlapping at room temperature, provided the rotational lines are described by a dispersion formula.[3] Emissivity estimates at elevated total pressures can be carried out without difficulty by using the tabulated numerical data presented previously (I). It is the purpose of the present analysis to provide useful approximate relations for emissivity calculations on diatomic gases at all temperatures for the special case when no overlapping occurs between collision-broadened rotational lines.

In Sec. II we present the basic relations for the engineering emissivity expressed as a sum of terms representing the contributions from individual vibration-rotation bands. Approximate relations are obtained on the assumption that the spectral half-widths for all rotational lines corresponding to the vibrational transitions $n \rightarrow n'$ are identical.

The results of extensive emissivity calculations as a function of temperature and optical density are presented in Sec. III for CO and HCl under conditions which will necessarily overestimate the correct values of the emissivity provided suitable numerical values have been used for integrated absorption and for rotational half-width.† For large values of the optical density, a treatment based on nonoverlapping rotational lines becomes a very poor approximation since, in particular, spectral emissivities greater than unity can be obtained. In order to avoid difficulties of this sort, it is convenient to utilize the known limiting emissivities (I).

In Sec. IV the calculated emissivities for CO are compared with empirical measurements. Possible causes for observed discrepancies are considered.

II. BASIC RELATIONS

The engineering emissivity E is defined by the relation

$$E=\int_0^\infty R_\omega d\omega \Big/ \int_0^\infty R_\omega{}^0 d\omega, \tag{1}$$

where R_ω is the radiant energy emitted per unit area per unit time by the emitter under study into a solid angle of 2π-steradians in the wave-number interval between ω and $\omega+d\omega$ and $R_\omega{}^0$ denotes the corresponding quantity for a blackbody. For uniformly distributed gaseous emitters

$$R_\omega=R_\omega{}^0[1-\exp(-P_\omega X)], \tag{2}$$

where P_ω denotes the spectral absorption coefficient. From Eqs. (1) and (2) it follows that

$$E=\int_0^\infty R_\omega{}^0[1-\exp(-P_\omega X)]d\omega/\sigma T^4, \tag{3}$$

where T is the temperature of the radiator and σ denotes the Stefan-Boltzmann constant.

* Supported by the ONR under Contract Nonr-220(03), NR 015 210. Presented, in part, before the Symposium on Molecular Structure and Spectroscopy, Columbus, June, 1952.

[1] S. S. Penner, J. Appl. Phys. **21**, 685 (1950); J. Appl. Mech. **18**, 53 (1951). Hereafter referred to as I.

[2] S. S. Penner and D. Weber, J. Appl. Phys. **22**, 1164 (1951). Hereafter referred to as II.

[3] Penner, Ostrander, and Tsien, J. Appl. Phys. **23**, 256 (1952). Hereafter referred to as III.

† The term rotational half-width is used in the present report to denote one-half of the wave-number range for which the spectral absorption coefficient exceeds one-half of the maximum value for a given rotational transition. Thus it corresponds to the quantity which is sometimes referred to as the semihalf-width.

For nonoverlapping rotational lines with a Lorentz distribution and of half-width $b_{j\to j'}{}^{n\to n'}$ for the rotational transition $j\to j'$ and for the vibrational transition $n\to n'$, it is readily shown that[4-6]

$$E=(\sigma T^4)^{-1}\sum_{j=0}^{\infty}\sum_{n=0}^{\infty}\sum_{n'>n}^{\infty}R^0(\omega_{j\to j'}{}^{n\to n'})\times[2\pi(b_{j\to j'}{}^{n\to n'})f(x_{j\to j'}{}^{n\to n'})], \quad (4)$$

where

$$f(x_{j\to j'}{}^{n\to n'})=x_{j\to j'}{}^{n\to n'}\exp(-x_{j\to j'}{}^{n\to n'})\times[J_0(ix_{j\to j'}{}^{n\to n'})-iJ_1(ix_{j\to j'}{}^{n\to n'})] \quad (5)$$

and

$$x_{j\to j'}{}^{n\to n'}=S_{j\to j'}{}^{n\to n'}X/2\pi b_{j\to j'}{}^{n\to n'}.$$

Here $S_{j\to j'}{}^{n\to n'}$ is the integrated absorption (expressed in cm^{-2} atoms^{-1}) of the rotational line associated with the rotational transition $j\to j'$ and the vibrational transition $n\to n'$, X is the optical density of the uniformly distributed emitters (expressed in cm atoms), and J_0 and J_1 denote Bessel functions of order zero and one, respectively. For optical densities at which the approximation of nonoverlapping between rotational lines applies, the values of n' which need to be considered normally are $n+1$ (fundamental vibration-rotation band) and $n+2$ (first overtone). The selection rules restrict the allowed values of j' to $j\pm1$ for molecules with P and R branches and to $j'=j$, $j\pm1$ for diatomic molecules with Q-branch (e.g., NO).

The values of $S_{j\to j'}{}^{n\to n'}$ are given, in good approximation, by the relations[6,7]

$$S_{j\to j-1}{}^{0\to1}=\alpha_F(\omega_{j\to j-1}{}^{0\to1}/\omega^*)[(j+\lambda)(j-\lambda)/j]\times[\exp-E(0,j)/kT]F_{j\to j-1}G_{j\to j-1}{}^{0\to1}Q^{-1}, \quad (6)$$

$$S_{j-1\to j}{}^{0\to1}=\alpha_F(\omega_{j-1\to j}{}^{0\to1}/\omega^*)[(j+\lambda)(j-\lambda)/j]\times[\exp-E(0,j-1)/kT]F_{j-1\to j}G_{j-1\to j}{}^{0\to1}Q^{-1}, \quad (7)$$

$$S_{j\to j}{}^{0\to1}=\alpha_F(\omega_{j\to j}{}^{0\to1}/\omega^*)[\lambda^2(2j+1)/j(j+1)]\times[\exp-E(0,j)/kT]G_{j\to j}{}^{0\to1}Q^{-1}, \quad (8)$$

$$S_{j\to j'}{}^{n\to n+1}/S_{j\to j'}{}^{0\to1}=(n+1)(\omega_{j\to j'}{}^{n\to n+1}/\omega_{j\to j'}{}^{0\to1})\times\{\exp[-E(n,j)+E(0,j)]/kT\}\times(G_{j\to j'}{}^{n\to n+1}/G_{j\to j'}{}^{0\to1}), \quad (9)$$

$$S_{j\to j'}{}^{0\to2}/S_{j\to j'}{}^{0\to1}=(\omega_{j\to j'}{}^{0\to2}/2\omega^*)(\omega^*/\omega_{j\to j'}{}^{0\to1})\times(G_{j\to j'}{}^{0\to2}/G_{j\to j'}{}^{0\to1})(\alpha_0/\alpha_F)\simeq(\alpha_0/\alpha_F)\times[1+\exp(-hc\omega^*/kT)], \quad (10)$$

$$S_{j\to j'}{}^{n\to n+2}/S_{j\to j'}{}^{0\to2}=[(n+2)(n+1)/2]\times(\omega_{j\to j'}{}^{n\to n+2}/\omega_{j\to j'}{}^{0\to2})\times\{\exp[-E(n,j)+E(0,j)]/kT\}\times(G_{j\to j'}{}^{n\to n+2}/G_{j\to j'}{}^{0\to2}), \quad (11)$$

$$\alpha_F=\sum_{j}^{\infty}\sum_{n}^{\infty}S_{j\to j'}{}^{n\to n+1}\simeq(\omega^*/\omega_e)(N_T\pi\epsilon^2/3\mu c), \quad (12)$$

$$\alpha_0=[1-\exp(-hc\omega^*/kT)]\times[1+\exp(-hc\omega^*/kT)]^{-1}\sum_{j}^{\infty}\sum_{n}^{\infty}S_{j\to j'}{}^{n\to n+2}=(N_Th\epsilon^2/24\pi c^3\omega_e\mu^2r_e{}^2)(\omega^*/\omega_e)(\rho_e''+a_1)^2.\ddagger \quad (13)$$

Here ω^* = wave number corresponding to the transition $j=0\to j=0$, $n=0\to n=1$; ω_e = wave number corresponding to an infinitesimal oscillation at the equilibrium internuclear distance; λ = quantum number measuring the electronic angular momentum about the internuclear axis; N_T = total number of molecules per unit volume; ϵ = rate of change of electric moment with internuclear distance; μ = reduced mass of the molecule; c = velocity of light; Q = complete partition function; $E(n,j)$ = energy of the nth vibrational and jth rotational level;

$$G_{j\to j'}{}^{n\to n'}=1-\exp(-hc\omega_{j\to j'}{}^{n\to n'}/kT);$$

$$F_{j\to j-1}=1+8\gamma j[1+(5\gamma j/4)-(3\gamma/4)];$$

$$F_{j-1\to j}=1-8\gamma j[1-(5\gamma j/4)-(3\gamma/4)],$$

where $\gamma=h/8\pi^2Ic\omega_e$ with I representing the equilibrium moment of inertia; r_e = equilibrium internuclear distance; $\rho_e''=p_e''/p_e'$, where p_e' and $p_e''/2$ represent, respectively, the coefficients multiplying $\xi=(r-r_e)/r_e$ and ξ^2 in the Taylor series expansion of the electric moment $p(\xi)$ about $\xi=0$; a_1 = mechanical anharmonicity constant, which is related to the usual spectroscopic constants through the equation $-\alpha_e=3\gamma B_e(a_1+1)$. The quantities α_F and α_0 must be determined empirically.

Let

$$E_{F1}=(\sigma T^4)^{-1}\sum_{j=0}^{\infty}R^0(\omega_{j\to j'}{}^{0\to1})\times[2\pi(b_{j\to j'}{}^{0\to1})f(x_{j\to j'}{}^{0\to1})] \quad (14)$$

represent the contribution to the total emissivity made by rotational lines belonging to the $n=0\to n=1$ vibrational transition. Equation (4) can evidently be written

[4] R. Ladenburg and F. Reiche, Ann. Physik **42**, 181 (1913).
[5] W. M. Elsasser, *Harvard Meteorological Studies No.* **6** (Blue Hill Observatory, 1942).
[6] S. S. Penner, J. Chem. Phys. **19**, 272, 1434 (1951).
[7] B. L. Crawford, Jr., and H. L. Dinsmore, J. Chem. Phys. **18**, 983, 1682 (1950).

‡ Equation (13) has been obtained from the analysis of Crawford and Dinsmore (reference 7). It is somewhat surprising that $\Sigma_j\Sigma_n S_{j\to j'}{}^{0\to2}$ has temperature-dependent terms for a given number of molecules, even in first order.

in the form

$$E=(\sigma T^4)^{-1}\sum_{j=0}^{\infty}R^0(\omega_{j\to j'}{}^{0\to1})[2\pi(b_{j\to j'}{}^{0\to1})f(x_{j\to j'}{}^{0\to1})]$$
$$\times\sum_{n=0}^{\infty}\sum_{n'>n}^{\infty}[R^0(\omega_{j\to j'}{}^{n\to n'})/R^0(\omega_{j\to j'}{}^{0\to1})]$$
$$\times[2\pi(b_{j\to j'}{}^{n\to n'})f(x_{j\to j'}{}^{n\to n'})/2\pi(b_{j\to j'}{}^{0\to1})$$
$$\times f(x_{j\to j'}{}^{0\to1})]. \quad (15)$$

Since E_{F1} can be calculated without difficulty (III), we shall express E in terms of E_{F1} by using Eq. (15) or a suitable modification of Eq. (15) successively for the vibrational transitions $0\to2$, $n\to n+1$, and $n\to n+2$. The calculations are simplified greatly by utilizing approximate expressions for the ratios appearing in the square brackets of Eq. (15).

The following relation follows obviously from the Planck blackbody distribution law:

$$R^0(\omega_{j\to j'}{}^{n\to n'})/R^0(\omega_{j\to j'}{}^{0\to1})=(\omega_{j\to j'}{}^{n\to n'}/\omega_{j\to j'}{}^{0\to1})^3$$
$$\times\{[\exp(u\omega_{j\to j'}{}^{0\to1}/\omega^*)]-1\}$$
$$\times\{[\exp(u\omega_{j\to j'}{}^{n\to n'}/\omega^*)]-1\}^{-1}, \quad (16)$$

where $u=hc\omega^*/kT$. A general expression cannot readily be obtained for the ratio $[2\pi(b_{j\to j'}{}^{n\to n'})f(x_{j\to j'}{}^{n\to n'})/2\pi(b_{j\to j'}{}^{0\to1})f(x_{j\to j'}{}^{0\to1})]$. However, use of appropriate asymptotic forms leads to the relations

$$2\pi(b_{j\to j'}{}^{n\to n'})f(x_{j\to j'}{}^{n\to n'})\simeq2(S_{j\to j'}{}^{n\to n'}b_{j\to j'}{}^{n\to n'}X)^{\frac{1}{2}}$$
for large values of $x_{j\to j'}{}^{n\to n'}$, (17a)

$$2\pi(b_{j\to j'}{}^{n\to n'})f(x_{j\to j'}{}^{n\to n'})\simeq S_{j\to j'}{}^{n\to n'}X$$
for small values of $x_{j\to j'}{}^{n\to n'}$. (17b)

It is well known that Eq. (17a) forms a useful approximation for values of $x_{j\to j'}{}^{n\to n'}$ as small as three.[4,5] It is evident that the computation of emissivities as a function of temperature will generally require the use of Eq. (17a) for n less than some small integer, whereas for larger values of n the approximation given in Eq. (17b) will be more reasonable. For what particular values of n and n' a change is to be effected will evidently depend on the numerical values of the optical density, integrated absorption, and rotational halfwidth.

We proceed now to use Eqs. (15) to (17b) in conjunction with Eqs. (6) to (13) to obtain explicit relations for the contributions to the total emissivity made by various vibration-rotation bands. Unfortunately the values of $b_{j\to j'}{}^{n\to n'}$ are usually not known whence we shall have to assume that $b_{j\to j'}{}^{n\to n+1}=b_{j\to j'}{}^{0\to1}=b_F$ and $b_{j\to j'}{}^{n\to n+2}=b_{j\to j'}{}^{0\to2}=b_0$.§ Since the condition of non-overlapping between rotational lines is not generally met for the Q branch and since, furthermore, NO is the only important diatomic molecule which exhibits a (weak) Q branch, further discussion will be restricted to the evaluation of E for diatomic molecules which have only P and R branches ($\lambda=0$).

A. Contributions to E made by the Vibrational Transition $n=0\to n=2$ (E_{01})

Since $\omega_{j\to j'}{}^{0\to2}\simeq2\omega_{j\to j'}{}^{0\to1}\simeq2\omega^*$, Eq. (16) becomes

$$R^0(\omega_{j\to j'}{}^{0\to2})/R^0(\omega_{j\to j'}{}^{0\to1})\simeq8[\exp(u)+1]^{-1}.\|$$

From Eqs. (10), (17a), and (17b) it is evident that

$$2\pi b_0 f(x_{j\to j'}{}^{0\to2})/2\pi b_F f(x_{j\to j'}{}^{0\to1})\simeq(\alpha_0/\alpha_F)[1+\exp(-u)]$$

for small values of $x_{j\to j'}{}^{0\to1}$, $x_{j\to j'}{}^{0\to2}$;

$$2\pi b_0 f(x_{j\to j'}{}^{0\to2})/2\pi b_F f(x_{j\to j'}{}^{0\to1})$$
$$\simeq(\alpha_0/\alpha_F)^{\frac{1}{2}}(b_0/b_F)^{\frac{1}{2}}[1+\exp(-u)]^{\frac{1}{2}}$$

for large values of $x_{j\to j'}{}^{0\to1}$, $x_{j\to j'}{}^{0\to2}$;

$$2\pi b_0 f(x_{j\to j'}{}^{0\to2})/2\pi b_F f(x_{j\to j'}{}^{0\to1})$$
$$\simeq\tfrac{1}{2}(\alpha_0/\alpha_F)(1/b_F)(S_{j\to j'}{}^{0\to1}b_F X)^{\frac{1}{2}}[1+\exp(-u)]$$

for large values of $x_{j\to j'}{}^{0\to1}$ and small values of $x_{j\to j'}{}^{0\to2}$.

We now obtain several special cases.

(a) If $x_{j\to j'}{}^{0\to1}$ and $x_{j\to j'}{}^{0\to2}$ are small, then

$$E_{01}\simeq8[\exp(-u)](\alpha_0/\alpha_F)E_{F1}. \quad (18a)$$

Equation (18a) would be expected to apply only to very weak radiators, at low optical densities, and at the largest pressures which are consistent with the idea that no overlapping occurs between rotational lines.

(b) If $x_{j\to j'}{}^{0\to1}$ and $x_{j\to j'}{}^{0\to2}$ are large, then

$$E_{01}\simeq8[\exp(-u)][1+\exp(-u)]^{-\frac{1}{2}}$$
$$\times(\alpha_0/\alpha_F)^{\frac{1}{2}}(b_0/b_F)^{\frac{1}{2}}E_{F1}. \quad (18b)$$

Equation (18b) would be expected to apply to intense radiators at moderate or large values of the optical density, particularly at low pressures where b_0 and b_F are small.

(c) If $x_{j\to j'}{}^{0\to1}$ is large and $x_{j\to j'}{}^{0\to2}$ is small, then

$$E_{01}\simeq8(\alpha_0/\alpha_F)(1/b_F)[\exp(-u)]b_F X(\sigma T^4)^{-1}$$
$$\times\sum_{j=0}^{\infty}R^0(\omega_{j\to j'}{}^{0\to1})S_{j\to j'}{}^{0\to1}.$$

But $R^0(\omega_{j\to j'}{}^{0\to1})\simeq R^0(\omega^*)$ and

$$\sum_{j=0}^{\infty}S_{j\to j'}{}^{0\to1}\equiv\alpha_{F1}\simeq\alpha_F[1-\exp(-u)]^2.^7$$

Hence,

$$E_{01}\simeq120u^5\pi^{-4}(\alpha_0/\alpha_F)[(\alpha_F/u)X/\omega^*]$$
$$\times[\exp(-2u)][1-\exp(-u)], \quad (18c)$$

§ This approximation is discussed more fully in paragraph IIG and in Appendix I.

‖ For large values of u (i.e., low temperatures) it is evident that $R^0(\omega_{j\to j'}{}^{0\to2})/R^0(\omega_{j\to j'}{}^{0\to1})\simeq8\exp(-u)$. For small values of u (i.e., high temperatures) it follows that $R^0(\omega_{j\to j'}{}^{0\to2})/R^0(\omega_{j\to j'}{}^{0\to1})\simeq4$.

where α_F/u is independent of temperature. Since α_0/α_F is generally very small compared to unity, Eq. (18c) will apply to many problems which are of practical interest.

B. Contributions to E made by the Vibrational Transitions $n\rightarrow n+1$ ($E_F=\sum_i E_{Fi}$).

Since $\omega_{j\rightarrow j'}{}^{n\rightarrow n+1}\simeq\omega_{j\rightarrow j'}{}^{0\rightarrow 1}\simeq\omega^*$, Eq. (16) becomes

$$R^0(\omega_{j\rightarrow j'}{}^{n\rightarrow n+1})/R^0(\omega_{j\rightarrow j'}{}^{0\rightarrow 1})\simeq 1.$$

From Eqs. (9), (17a), and (17b) it can be seen that

$$2\pi b_F f(x_{j\rightarrow j'}{}^{n\rightarrow n+1})/2\pi b_F f(x_{j\rightarrow j'}{}^{0\rightarrow 1})=(n+1)\exp(-nu)$$

for small values of $x_{j\rightarrow j'}{}^{0\rightarrow 1}$, $x_{j\rightarrow j'}{}^{n\rightarrow n+1}$;

$$2\pi b_F f(x_{j\rightarrow j'}{}^{n\rightarrow n+1})/2\pi b_F f(x_{j\rightarrow j'}{}^{0\rightarrow 1})=(n+1)^{\frac{1}{2}}\exp(-nu/2)$$

for large values of $x_{j\rightarrow j'}{}^{0\rightarrow 1}$, $x_{j\rightarrow j'}{}^{n\rightarrow n+1}$;

$$2\pi b_F f(x_{j\rightarrow j'}{}^{n\rightarrow n+1})/2\pi b_F f(x_{j\rightarrow j'}{}^{0\rightarrow 1})=\tfrac{1}{2}(n+1)\exp(-nu)(S_{j\rightarrow j'}{}^{0\rightarrow 1}b_F X)^{\frac{1}{2}}(b_F)^{-1}$$

for large values of $x_{j\rightarrow j'}{}^{0\rightarrow 1}$ and small values of $x_{j\rightarrow j'}{}^{n\rightarrow n+1}$. It is again of interest to examine several special cases.

(a) If $x_{j\rightarrow j'}{}^{0\rightarrow 1}$ and $x_{j\rightarrow j'}{}^{n\rightarrow n+1}$ are small, then

$$E_F=E_{F1}[1-\exp(-u)]^{-2}. \tag{19a}$$

Equation (19a) applies under the same conditions as Eq. (18a).

(b) If $x_{j\rightarrow j'}{}^{0\rightarrow 1}$ and $x_{j\rightarrow j'}{}^{n\rightarrow n+1}$ are large, then

$$E_F=\left\{\sum_{n=0}^{\infty}(n+1)^{\frac{1}{2}}[\exp(-nu/2)]\right\}E_{F1}. \tag{19b}$$

It is evident that Eq. (19b) cannot apply in any real problem since $x_{j\rightarrow j'}{}^{n\rightarrow n+1}$ must become small for sufficiently large values of n.

(c) If $x_{j\rightarrow j'}{}^{0\rightarrow 1}$ is large and $x_{j\rightarrow j'}{}^{n\rightarrow n+1}$ is large for $n<K$ but $x_{j\rightarrow j'}{}^{n\rightarrow n+1}$ is small for $n\gtrsim K$, then

$$E_F=E_{F1}\left\{\sum_{n=0}^{K-1}(n+1)^{\frac{1}{2}}[\exp(-nu/2)]\right\}+\left\{\sum_{n=K}^{\infty}(n+1)[\exp(-nu)]\right\}X(\sigma T^4)^{-1}\times\sum_{j=0}^{\infty}R^0(\omega_{j\rightarrow j'}{}^{0\rightarrow 1})S_{j\rightarrow j'}{}^{0\rightarrow 1}.$$

But $R^0(\omega_{j\rightarrow j'}{}^{0\rightarrow 1})\simeq R^0(\omega^*)$, whence

$$E_F\simeq E_{F1}\left\{\sum_{n=0}^{K-1}(n+1)^{\frac{1}{2}}[\exp(-nu/2)]\right\}+15u^5\pi^{-4}[(\alpha_F/u)X/\omega^*]\times\exp(-u)[1-\exp(-u)]\times\left\{\sum_{n=K}^{\infty}(n+1)[\exp(-nu)]\right\}, \tag{19c}$$

where the sum over n may be evaluated conveniently by noting that

$$\sum_{n=K}^{\infty}(n+1)[\exp(-nu)]=[1-\exp(-u)]^{-2}-\sum_{n=0}^{K-1}(n+1)[\exp(-nu)].$$

Equation (19c) is a useful expression for E_F for most problems which arise in practice. Depending upon the numerical values of $S_{j\rightarrow j'}{}^{0\rightarrow 1}$, X, b_F, and the temperature, the appropriate value for K will be an integer somewhere between 1 and 5.

C. Contributions to E made by the Vibrational Transitions $n\rightarrow n+2$ ($E_0=\sum_i E_{0i}$).

For the evaluation of E_0 it is convenient to replace Eq. (15) by the relation

$$E_0=(\sigma T^4)^{-1}\sum_{j=0}^{\infty}R^0(\omega_{j\rightarrow j'}{}^{0\rightarrow 2})[2\pi b_0 f(x_{j\rightarrow j'}{}^{0\rightarrow 2})]\times\sum_{n=0}^{\infty}[R^0(\omega_{j\rightarrow j'}{}^{n\rightarrow n+2})/R^0(\omega_{j\rightarrow j'}{}^{0\rightarrow 2})]\times[2\pi b_0 f(x_{j\rightarrow j'}{}^{n\rightarrow n+2})/2\pi b_0 f(x_{j\rightarrow j'}{}^{0\rightarrow 2})]. \tag{20}$$

By proceeding from Eq. (16) we obtain the result

$$R^0(\omega_{j\rightarrow j'}{}^{n\rightarrow n+2})/R^0(\omega_{j\rightarrow j'}{}^{0\rightarrow 2})\simeq 1$$

since

$$\omega_{j\rightarrow j'}{}^{n\rightarrow n+2}\simeq\omega_{j\rightarrow j'}{}^{0\rightarrow 2}.$$

Equations (11), (17a), and (17b) lead to the relations

$$2\pi b_0 f(x_{j\rightarrow j'}{}^{n\rightarrow n+2})/2\pi b_0 f(x_{j\rightarrow j'}{}^{0\rightarrow 2})=[(n+2)(n+1)/2][\exp(-nu)]$$

for small values of $x_{j\rightarrow j'}{}^{0\rightarrow 2}$ and $x_{j\rightarrow j'}{}^{n\rightarrow n+2}$;

$$2\pi b_0 f(x_{j\rightarrow j'}{}^{n\rightarrow n+2})/2\pi b_0 f(x_{j\rightarrow j'}{}^{0\rightarrow 2})=[(n+2)(n+1)/2]^{\frac{1}{2}}[\exp(-nu/2)]$$

for large values of $x_{j\rightarrow j'}{}^{0\rightarrow 2}$ and $x_{j\rightarrow j'}{}^{n\rightarrow n+2}$;

$$2\pi b_0 f(x_{j\rightarrow j'}{}^{n\rightarrow n+2})/2\pi b_0 f(x_{j\rightarrow j'}{}^{0\rightarrow 2})=\tfrac{1}{2}[(n+2)(n+1)/2]\exp(-nu)(S_{j\rightarrow j'}{}^{0\rightarrow 2}b_0 X)^{\frac{1}{2}}(b_0)^{-1}$$

for large values of $x_{j\rightarrow j'}{}^{0\rightarrow 2}$ and small values of $x_{j\rightarrow j'}{}^{n\rightarrow n+2}$.

Thus we obtain for E_0 the following approximate relations:

(a) For small values of $x_{j\rightarrow j'}{}^{0\rightarrow 2}$ and $x_{j\rightarrow j'}{}^{n\rightarrow n+2}$:

$$E_0=\left\{\sum_{n=0}^{\infty}[(n+2)(n+1)/2][\exp(-nu)]\right\}E_{01}=[1-\exp(-u)]^3E_{01}. \tag{20a}$$

Equation (20a) constitutes a useful approximation for

moderate or small values of the optical density for most diatomic molecules.

(b) For large values of $x_{j\to j'}{}^{0\to 2}$ and $x_{j\to j'}{}^{n\to n+2}$

$$E_0=\left\{\sum_{n=0}^{\infty}[(n+2)(n+1)/2]^{\frac{1}{2}}[\exp(-nu/2)]\right\}E_{01}. \quad (20b)$$

Equation (20b) is of no practical interest since $x_{j\to j'}{}^{n\to n+2}$ must become very small at all temperatures for sufficiently large values of n.

(c) For large values of $x_{j\to j'}{}^{0\to 2}$ and $x_{j\to j'}{}^{n\to n+2}$ for $n<K$ and small values of $x_{j\to j'}{}^{n\to n+2}$ for $n\geqslant K$,

$$E_0=\left\{\sum_{n=0}^{K-1}[(n+2)(n+1)/2]^{\frac{1}{2}}[\exp(-nu/2)]\right\}E_{01}$$
$$+120(u/\pi)^4(\alpha_{01}X/\omega^*)\{[\exp(2u)]-1\}^{-1}$$
$$\times\left\{\sum_{n=K}^{\infty}[(n+2)(n+1)/2][\exp(-nu)]\right\},$$

where[7]

$$\alpha_{01}=\sum_{j=0}^{\infty}S_{j\to j'}{}^{0\to 2}\simeq\alpha_0[1-\exp(-u)][1-\exp(-2u)].$$

Hence,

$$E_0=\left\{\sum_{n=0}^{K-1}[(n+2)(n+1)/2]^{\frac{1}{2}}[\exp(-nu/2)]\right\}E_{01}$$
$$+120u^5\pi^{-4}(\alpha_0/\alpha_F)[(\alpha_F/u)X/\omega^*]$$
$$\times[\exp(-2u)][1-\exp(-u)]$$
$$\times\left\{\sum_{n=K}^{\infty}[(n+2)(n+1)/2][\exp(-nu)]\right\}. \quad (20c)$$

D. Contributions to E made by Vibrational Transitions $n\to n'$, $n'>n+2$.

The total contributions to radiant heat transfer from harmonics above the first overtone can be calculated by an obvious extension of the methods described above by using appropriate relations obtainable from the analysis of Crawford and Dinsmore.[7]

E. The Total Emissivity E

Suitable approximate expressions for the total engineering emissivity E can be obtained by combining the results given in the preceding paragraphs. A representative case, which may be considered to be a reasonable approximation for CO at small optical densities and for temperatures between 300°K and 1500°K, can be obtained as follows. The quantities $x_{j\to j'}{}^{0\to 1}$ and $x_{j\to j'}{}^{0\to 2}$ are considered to be large, whence it follows that E_{01} is given by Eq. (18b). Equation (19c) applies for E_F with $K=3$ and Eq. (20c) applies for E_0 with $K=1$. Thus,

$$E=E_{F1}\left\{\sum_{n=0}^{2}(n+1)^{\frac{1}{2}}[\exp(-nu/2)]+8[\exp(-u)]\right.$$
$$\left.\times[1+\exp(-u)]^{-\frac{1}{2}}(\alpha_0/\alpha_F)^{\frac{1}{2}}(b_0/b_F)^{\frac{1}{2}}\right\}$$
$$+15u^5\pi^{-4}[(\alpha_F/u)X/\omega^*][\exp(-u)]$$
$$\times[1-\exp(-u)]\{\{[1-\exp(-u)]^{-2}$$
$$-\sum_{n=0}^{2}(n+1)[\exp(-nu)]\}+8(\alpha_0/\alpha_F)$$
$$\times[\exp(-u)]\{[1-\exp(-u)]^{-3}-1\}\}. \quad (21)$$

Equation (21) has been written in such a way that E is given directly in terms of dimensionless parameters which depend only on the molecular constants and on E_{F1}.

An approximate expression for E_{F1} can be obtained very simply from Eq. (14) by utilizing the asymptotic form given in Eq. (17a) with the integrated absorption per rotational line given by Eqs. (6) and (7). Although the relation which is obtained by making such approximations as $R^0(\omega_{j\to j'}{}^{0\to 1})\simeq R^0(\omega^*)$ is entirely adequate for the present purposes, we shall not utilize the result obtained in this manner since a more accurate evaluation of E_{F1} from Eq. (14) has been carried out by Tsien (III). The resulting expression for E_{F1} is

$$E_{F1}=30u^5\pi^{-4}[\exp(-u)][(\gamma b_F/\omega^*)(\alpha_F/u)(X/\omega^*)]^{\frac{1}{2}}$$
$$\times f(\gamma,\delta,x^*,u)\{1.225(2/\gamma u)^{\frac{1}{2}}$$
$$\times[1+(3\gamma/2u)g(\gamma,\delta,u)]-0.417\}. \quad (22)$$

Here $f(\gamma,\delta,x^*,u)$ and $[1+(3\gamma/2u)g(\gamma,\delta,u)]$ are functions which are defined elsewhere but are generally close to unity (Sec. III for details).

The expression given in Eq. (21) would be expected to overestimate¶ the actual value of the engineering emissivity (a) because partial overlapping may occur between rotational lines and (b) because the asymptotic forms for $f(x)$ overestimate $f(x)$ for every value of x. Since the spectral emissivity of uniformly heated gases cannot exceed the spectral emissivity of a blackbody at the same temperature, Eq. (21) will evidently lead to invalid results for excessively large values of X. In this connection it will be useful to refer to the limiting emissivities computed in I.

We have developed two methods for the theoretical calculation of E, both of which overestimate the true value of E. For small and moderate values of X, Eq. (21) or a suitable modification of Eq. (21) represents a useful approximation provided the rotational spacing

¶ Equation (21) will yield values for E which are too small at temperatures which are so high that Doppler-broadening becomes important. For spectral lines which are best described by combined Doppler- and collision-broadening the present treatment will require extensive revision. See Sec. IV in this connection.

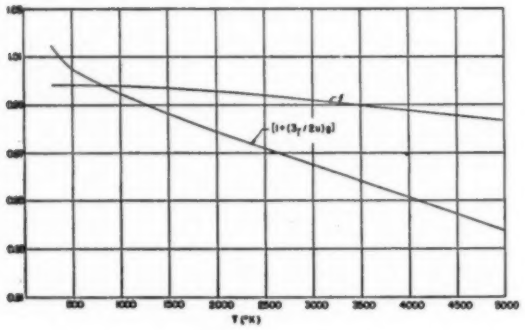


FIG. 1. The quantities f and $[1+(3\gamma/2u)g]$ as a function of temperature for CO.

is large compared to the collision half-width. For large values of X or at elevated pressures the use of average absorption coefficients is generally preferable (I). When very precise estimates are desired for the emissivity, it is possible to perform numerical calculations by evaluating the integral appearing in Eq. (3) directly. This procedure is indicated for partial overlapping between rotational lines and is facilitated by the use of computing machines.

F. Limits of Validity of the Theoretical Expression for E

The limits of validity of the relations derived in the preceding paragraphs for nonoverlapping rotational lines are defined by (a) the approximations involved in the use of asymptotic limiting forms and (b) the validity of the assumptions that

$$b_{j\to j'}{}^{n\to n+1}=b_{j\to j'}{}^{0\to 1}=b_F \quad \text{and} \quad b_{j\to j'}{}^{n\to n+2}=b_{j\to j'}{}^{0\to 2}=b_0.$$

The use of asymptotic limiting forms, in conjunction with the assumption that no overlapping occurs between rotational lines, will necessarily maximize the calculated values of E. This statement remains true independently of the choice of a particular one of the asymptotic forms given in Eqs. (17a) and (17b). Thus both x and $(2x/\pi)^{\frac{1}{2}}$ are larger than $f(x)$ for every value of x with $f(x)=x$ representing a better approximation for small values of x and $f(x)=(2x/\pi)^{\frac{1}{2}}$ being more adequate for large values of x. In general, the relation $f(x)=x$ is to be preferred for $x\leqslant(2/\pi)$ and the relation $f(x)=(2x/\pi)^{\frac{1}{2}}$ for $x\geqslant(2/\pi)$. Thus we arrive at the important conclusion that we have obtained reasonable upper limits for the emissivity as long as the line width and line shape are described correctly.

The effect on E of the assumed line shape and line width cannot be estimated in a satisfactory manner at the present time. Since we have neglected Doppler broadening entirely, the calculated values of E will be relatively too low as the temperature is raised. Representative calculations show, however, that at atmospheric pressure (and hence also at higher pressures) the Doppler half-width is generally small compared to the collision half-width at temperatures below about 2000°K (Sec. IV). Therefore, it appears unlikely that the calculated values of E will be reduced significantly, in the temperature range for which experimental data are available, merely because of neglect of the contributions of Doppler broadening to the spectral line shape and half-width. On the other hand, it is entirely possible that relatively large errors are introduced by the approximations $b_{j\to j'}{}^{n\to n+1}=b_F$ and $b_{j\to j'}{}^{n\to n+2}=b_0$. Although semiquantitative considerations of the type given in Appendix I suggest that the effective collision diameter of diatomic molecules is not a sensitive function of vibrational quantum number, it is hardly justifiable to apply this conclusion to the optical collision diameter. Thus we are forced to note that the calculated values of E are unreliable at temperatures above about 500°K where radiant heat transfer from excited vibrational energy states begins to play an important role. The theoretical relations obtained in the present discussion can be modified, without difficulty, to allow for a more adequate description of line width if and when the needed experimental and/or theoretical data become available.

III. EMISSIVITY CALCULATIONS FOR NONOVERLAPPING ROTATIONAL LINES FOR CO AND HCl**

Equation (21) will now be used for representative emissivity calculations on CO and HCl. First we shall compute E_{F1} from Eq. (22). Next the contributions to radiant heat transfer arising from the transitions $n=1\to n'=2$, $n=2\to n'=3$, and $n=0\to n'=2$ are evaluated. Finally the total emissivity E is determined according to Eq. (21).

A. Calculation of E_{F1}

The partial emissivity E_{F1} is determined by Eq. (22). Explicit expressions for f and $[1+(3\gamma/2u)g]$ have been given previously in III. Plots of these quantities are shown as a function of temperature for CO and HCl in Figs. 1 and 2, respectively. An adequate approximation

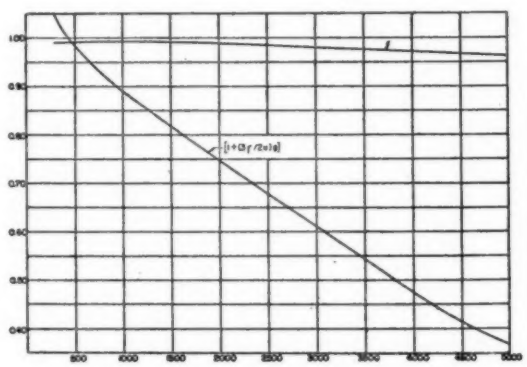


FIG. 2. The quantities f and $[1+(3\gamma/2u)g]$ as a function of temperature for HCl.

** Numerical calculations were carried out by Mr. E. K. Björnerud and Mrs. B. MacDonald.

to E_{F1} for CO at all temperatures is evidently given by the relation (III)

$$E_{F1}\simeq 0.6345u^5[\exp(-u)](\gamma u)^{-\frac{1}{2}} \times[(\gamma b/\omega^*)(X/\omega^*)(\alpha_F/u)]. \quad (22a)$$

The quantity E_{F1} has been computed as a function of temperature, according to Eq. (22) for CO and HCl. The numerical values were obtained by using the well-known result that the Lorentz half-width varies as $T^{-\frac{1}{2}}$ at constant pressure. For the conditions of interest in the present study it was assumed that

$$b\equiv b_c=b_F=b_0=0.076\,(T/300)^{-\frac{1}{2}}\ \text{cm}^{-1}\ \text{for CO}^{8}$$

and

$$b\equiv b_c=b_F=b_0=0.10\,(T/300)^{-\frac{1}{2}}\ \text{cm}^{-1}\ \text{for HCl},^{9}$$

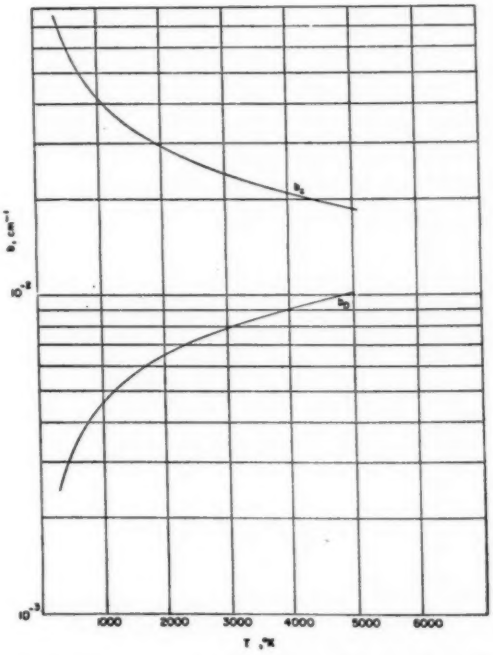


FIG. 3. Doppler half-width (b_D) and collision half-width (b_c) as a function of temperature for CO.

i.e., the spectral half-width b was set equal to the collision half-width b_c. It appears likely that Doppler-broadening will not make a significant contribution to E_{F1} at temperatures below about 2000°K. The Doppler half-widths (b_D) and collision half-widths (b_c) are plotted as a function of temperature in Figs. 3 and 4 for CO and HCl, respectively.

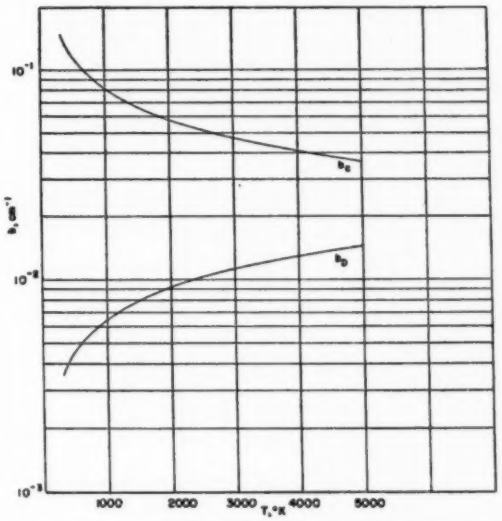


FIG. 4. Doppler half-width (b_D) and collision half-width (b_c) as a function of temperature for HCl.

The basic spectroscopic constants which were used for the numerical calculations are summarized in Table I. Numerical values for E_{F1}, as a function of temperature and optical density, have been calculated for CO and for HCl. The contributions to E made by the transitions $n=1\rightarrow n'=2$ (E_{F2}), $n=2\rightarrow n'=3$ (E_{F3}), $n=0\rightarrow n'=2$ (E_{01}), and $n\rightarrow n+1$ for $n>2$ ($E_F-E_{F1}-E_{F2}-E_{F3}$), $n\rightarrow n+2$ for $n>0$ (E_0-E_{01}) have also been determined. The quantities E_{F1}, E_{F2}, E_{F3}, $E_F-\sum_{i=1}^{3}E_{Fi}$, E_{01}, and E_0-E_{01} are plotted as a function of temperature for CO at optical densities of 0.30, 1.00, 2.00, and 10.00 cm atmos in Figs. 5 to 8, respectively. The values of

TABLE I. Basic spectroscopic constants for CO[a] and HCl.[b]

Spectroscopic constant	CO	HCl
ω_e, cm^{-1}	2169.81	2989.74
B_e, cm^{-1}	1.93130	10.5909
α_e	0.017543	0.3019
D_e	5.95×10^{-6}	5.32×10^{-4}
$x_e\omega_e=x^*\omega^*$	13.284	52.05
$y_e\omega_e$	0.010	0.056
ω^*, cm^{-1}	2143.14	2885.70
δ	9.0835×10^{-3}	2.85056×10^{-2}
x^*	6.1981×10^{-3}	1.8027×10^{-2}
$\gamma=B_e/\omega_e$	9.0111×10^{-4}	3.5424×10^{-3}
$uT=hc\omega^*/k$	3082.98	4150.99
α_0/u, cm^{-2}–atmos^{-1}	0.161[c]	0.264[9]
α_F/u, cm^{-2}–atmos^{-1}	22.95[c]	10.73[9]

$$E(n,j)=\omega_e(n+\tfrac{1}{2})-\omega_e x_e(n+\tfrac{1}{2})^2+\omega_e y_e(n+\tfrac{1}{2})^3 +B_e j(j+1)-\alpha_e(n+\tfrac{1}{2})j(j+1)-D_e j^2(j+1)^2$$

[a] Plyler, Benedict, and Silverman, J. Chem. Phys. **20**, 175 (1952).
[b] G. Herzberg, *Spectra of Diatomic Molecules* (D. Van Nostrand Company, Inc., New York, 1950), p. 534.
[c] S. S. Penner and D. Weber, J. Chem. Phys. **19**, 807, 817, 974 (1951).

[8] S. S. Penner and D. Weber, J. Chem. Phys. **19**, 1351, 1361 (1951). *Note added in proof:* The value $b_0=0.076$ cm^{-1} atmos^{-1} at room temperature is based on an extrapolation using unpublished measurements of b_0 as a function of experimental slit width. W. S. Benedict and S. Silverman have studied the first overtone of CO with greatly improved resolution and have reported $b_0=0.095$ cm^{-1} atmos^{-1} for self-broadening (paper presented at the Symposium on Molecular Structure and Spectroscopy, Columbus, June, 1952). The calculated emissivities are relatively insensitive to the numerical value used for b_0. Thus, changing b_0 by a factor of two will generally change E by less than 10 percent.

[9] S. S. Penner and D. Weber, paper presented at the *Symposium on Molecular Structure and Spectroscopy*, Ohio State University, Columbus, 1951.

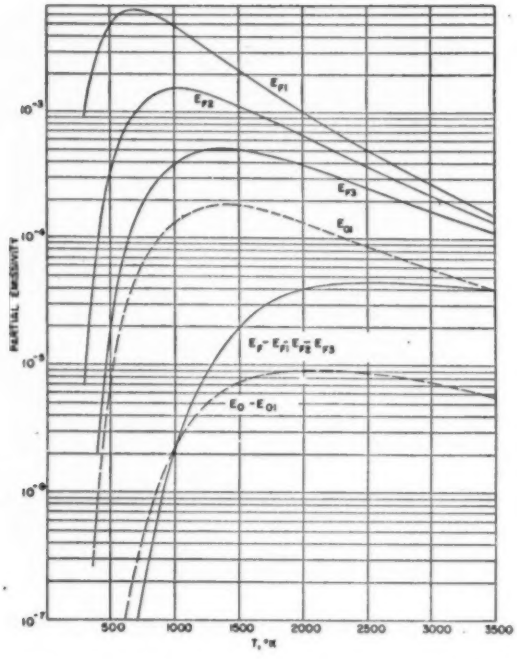


FIG. 5. Partial emissivities as a function of temperature for CO at $X=0.30$ cm atmos.

$E_{F1}, E_{F2}, E_{F3}, E_F - \sum_{i=1}^{3} E_{Fi}, E_{01}$, and $E_0 - E_{01}$ are plotted as a function of temperature for HCl at optical densities of 1, 5, 20, 50, and 100 cm atmos in Figs. 9 to 13, respectively.

The total emissivities E are plotted as a function of temperature for CO at optical densities of 0.30, 1.00, 2.00, and 10.00 cm atmos in Figs. 14 to 17, respectively. Also shown in Figs. 14 to 17 are experimentally observed

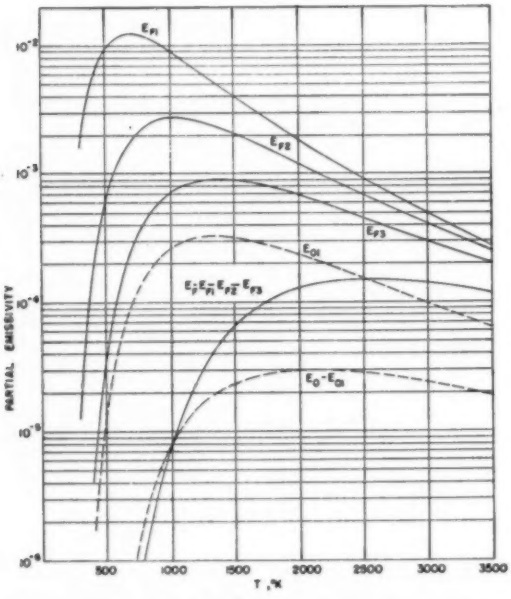


FIG. 6. Partial emissivities as a function of temperature for CO at $X=1.0$ cm atmos.

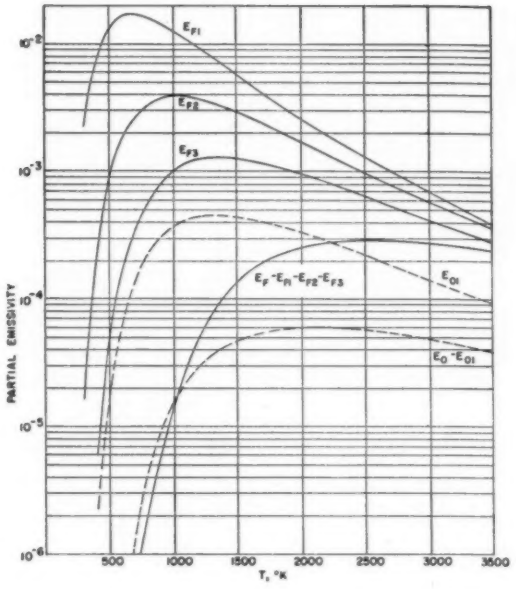


FIG. 7. Partial emissivities as a function of temperature for CO at $X=2.0$ cm atmos.

emissivities reported by Ullrich,[10] limiting emissivities E_L computed by the method described in I and the quantity $E_{F1}(T/300)$. The total emissivities E are plotted as a function of temperature for HCl at optical densities of 1, 5, 20, 50, and 100 cm-atmos in Figs. 18 to 22, respectively, together with the corresponding values of the limiting emissivity E_L.

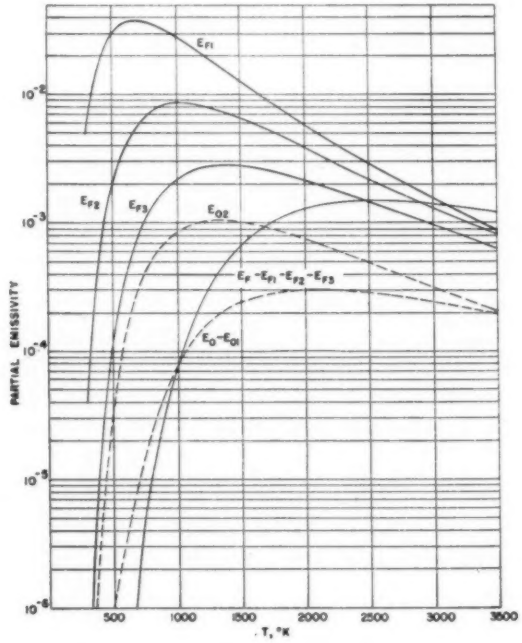


FIG. 8. Partial emissivities as a function of temperature for CO at $X=10$ cm atmos.

[10] W. Ullrich, Doctor of Science thesis, M.I.T., Cambridge, 1935.

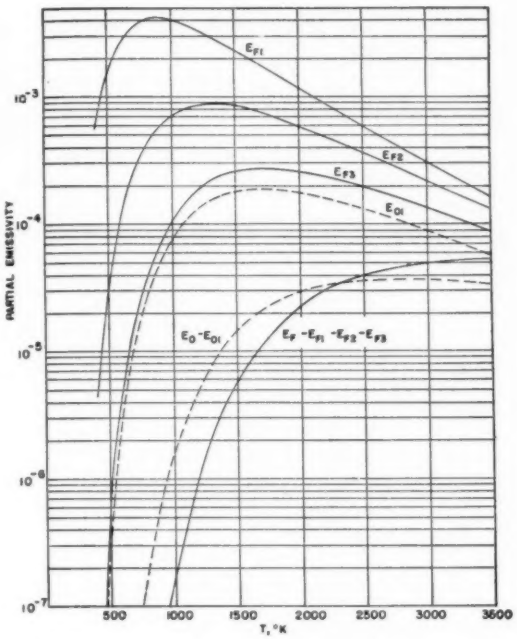


FIG. 9. Partial emissivities as a function of temperature for HCl at $X=1.0$ cm atmos.

IV. DISCUSSION OF RESULTS

The principal source of error in the present calculations is probably the result of inadequate description of line widths and line shape. In this connection it is of interest to note that the more adequate theories of line broadening developed by Weisskopf,[11] Lindholm,[12] Van Vleck and Weisskopf,[13] Van Vleck and Margenau,[14] Margenau,[15] and others[16] all lead to a decrease of collision half-width with temperature, which does not differ greatly from the assumed proportionality with $T^{-\frac{1}{2}}$.

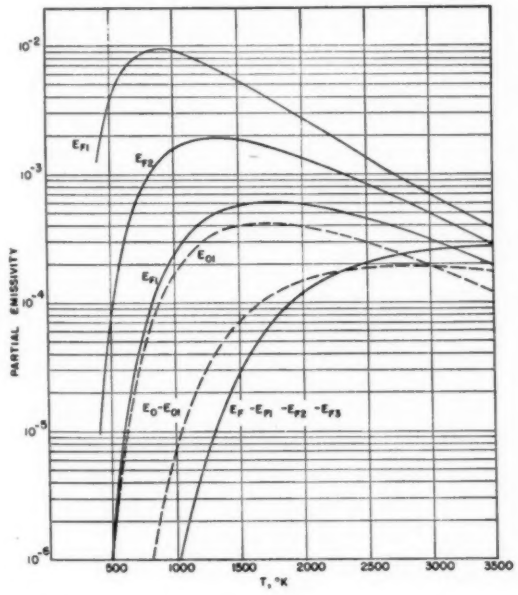


FIG. 10. Partial emissivities as a function of temperature for HCl at $X=5$ cm atmos.

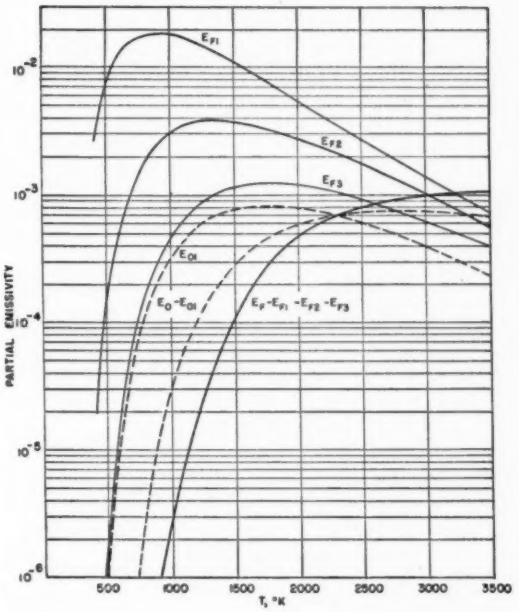


FIG. 11. Partial emissivities as a function of temperature for HCl at $X=20$ cm atmos.

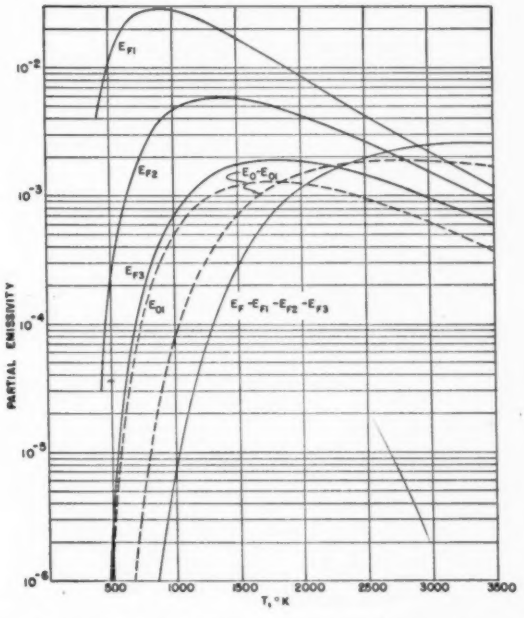


FIG. 12. Partial emissivities as a function of temperature for HCl at $X=50$ cm atmos.

[11] V. F. Weisskopf, Physik. Z. **34**, 1 (1933).
[12] E. Lindholm, dissertation, Uppsala, 1942.
[13] J. H. Van Vleck and V. F. Weisskopf, Revs. Modern Phys. **17**, 227 (1945).
[14] J. H. Van Vleck and H. Margenau, Phys. Rev. **76**, 1211 (1949). Van Vleck and Margenau have demonstrated the equivalence of line shapes in absorption and emission, a result which has been tacitly assumed throughout the present analysis.
[15] H. Margenau, Phys. Rev. **82**, 156 (1951).
[16] P. W. Anderson, Phys. Rev. **76**, 647 (1949).

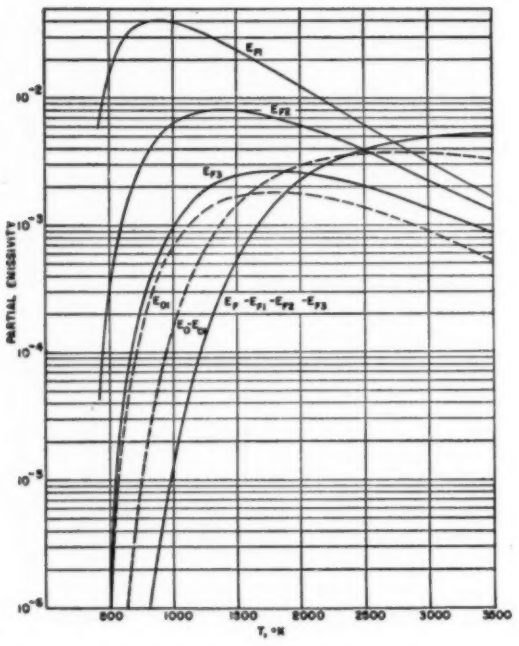


FIG. 13. Partial emissivities as a function of temperature for HCl at $X=100$ cm atmos.

In order to estimate the effect of Doppler broadening on the emissivity, representative numerical calculations of intensities for combined collision- and Doppler-broadening have been carried out. These calculations show that the effect of Doppler broadening on the calculated emissivities of CO must be negligibly small at temperatures below about 3000°K.

For HCl it is well known that the Lorentz collision formula does not represent experimentally determined transmission measurements adequately.[9,12] For this reason the emissivities calculated for nonoverlapping rotational lines should not be considered as being reliable to within much more than a factor of about two.

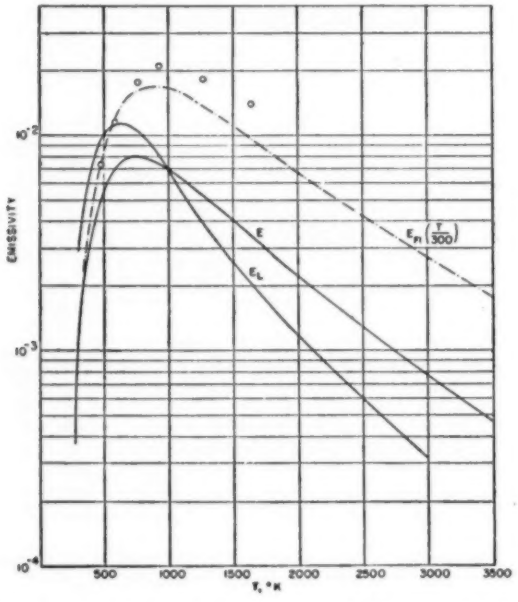


FIG. 14. Calculated and observed emissivities as a function of temperature for CO at $X=0.30$ cm atmos (emissivities based on treatment for nonoverlapping rotational lines: E; limiting emissivities: E_L; empirical relation: $E_{F1}(T/300)$; data obtained from empirical measurements: o).

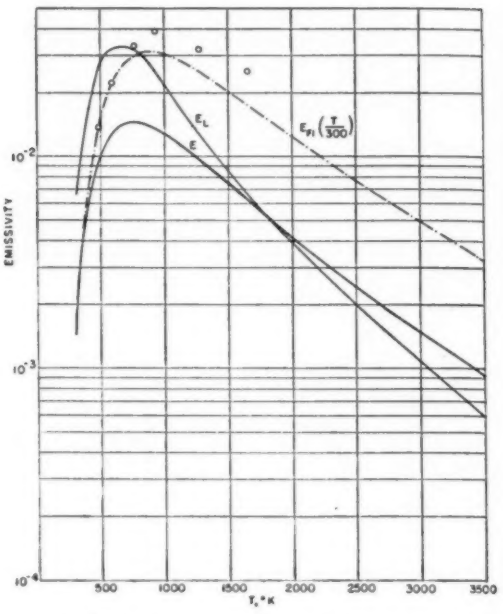


FIG. 15. Calculated and observed emissivities as a function of temperature for CO at $X=1.0$ cm atmos (emissivities based on treatment for nonoverlapping rotational lines: E; limiting emissivities: E_L; empirical relation: $E_{F1}(T/300)$; data obtained from empirical measurements: o).

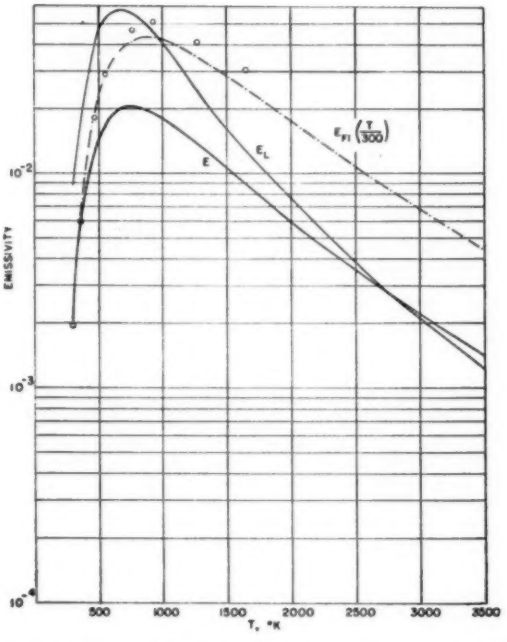


FIG. 16. Calculated and observed emissivities as a function of temperature for CO at $X=2.0$ cm atmos (emissivities based on treatment for nonoverlapping rotational lines: E; limiting emissivities: E_L; empirical relation: $E_{F1}(T/300)$; data obtained from empirical measurements: o).

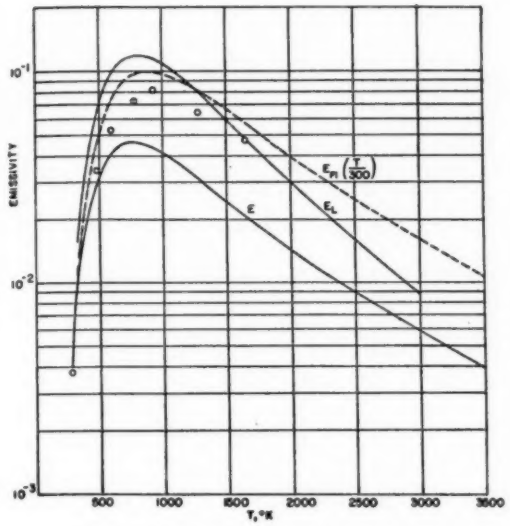


FIG. 17. Calculated and observed emissivities as a function of temperature for CO at $X=10.0$ cm atmos (emissivities based on treatment for nonoverlapping rotational lines: E; limiting emissivities: E_L; empirical relation: $E_{F1}(T/300)$; data obtained from empirical measurements: o).

To summarize, the calculated emissivities represent upper limits except in so far as the half-widths of excited vibrational energy states may have been underestimated. The optical densities for which numerical calculations have been performed were chosen in such a way that the asymptotic expressions used and the neglect of overlapping between rotational lines constitute very reasonable approximations. In fact, unless the collision half-width varies with temperature and vibrational quantum number in an unpredictable manner, it would be expected that the calculated emissivities are in error by less than about 20 percent.

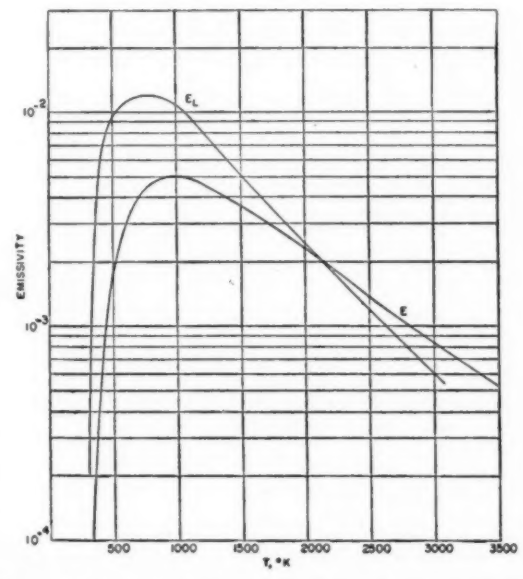


FIG. 18. Total emissivities as a function of temperature for HCl at $X=1$ cm atmos (emissivities based on treatment for nonoverlapping rotational lines: E; limiting emissivities: E_L).

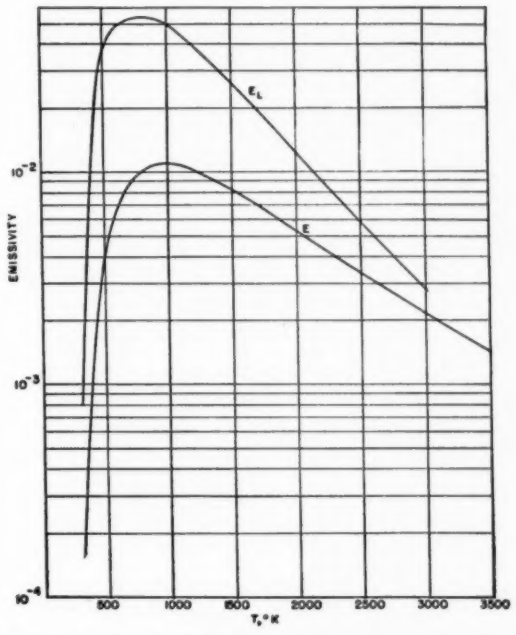


FIG. 19. Total emissivities as a function of temperature for HCl at $X=5$ cm atmos (emissivities based on treatment for nonoverlapping rotational lines: E; limiting emissivities: E_L).

Reference to Figs. 14 to 17 permits a direct comparison between emissivity values calculated from spectroscopic data and empirically observed emissivities for CO. It is apparent that the experimentally determined emissivities are represented quite well by the function $E_{F1}(T/300)$ for which no theoretical justification can be given. According to Ullrich,[10] the empirically determined values of the emissivities have a probable error of not over 20 percent for temperatures between about 500 and 1300°K if $0.5 \leqslant X \leqslant 5$ cm atmos. However, the

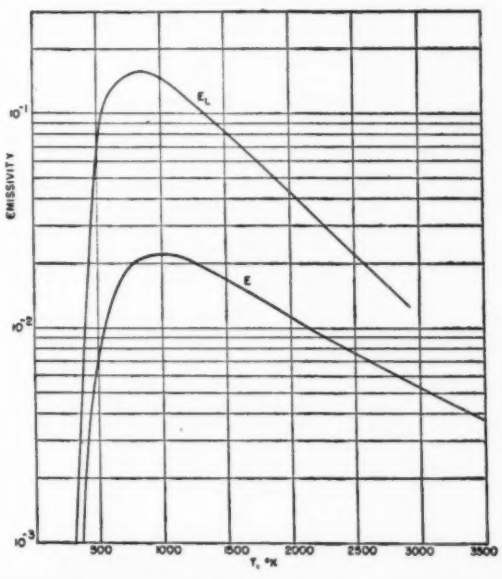


FIG. 20. Total emissivities as a function of temperature for HCl at $X=20$ cm atmos (emissivities based on treatment for nonoverlapping rotational lines: E; limiting emissivities: E_L).

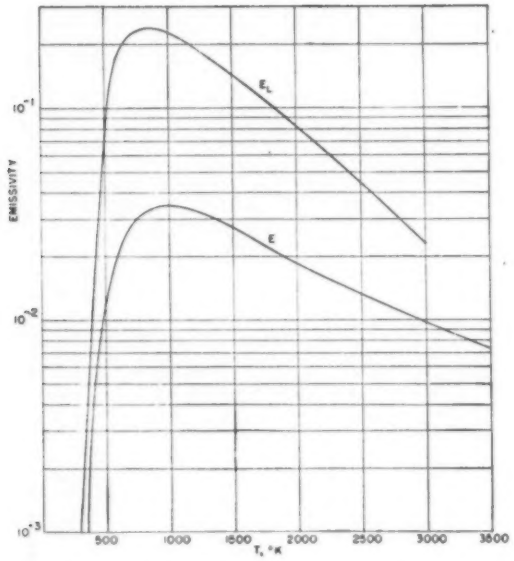


FIG. 21. Total emissivities as a function of temperature for HCl at $X=50$ cm atmos (emissivities based on treatment for non-overlapping rotational lines: E; limiting emissivities: E_L).

discrepancies between calculated and observed emissivities are outside of the probable limits of error of either investigation, except for the reservations noted previously with regard to dependence of line width on vibrational excitation. The result that the calculated emissivities deviate more from the empirical measurements the higher the temperature is consistent with the idea that the optical collision diameter increases rapidly with vibrational excitation. In this connection it may be of interest to note also that Ullrich had increasing difficulties with his experimental studies at elevated temperatures because of diffusion of air into the high temperature chamber containing the carbon monoxide.[10] It is evident that the apparent emissivities will be much too large if only a trace of CO_2 is formed in the reaction vessel because of the very large value for the integrated absorption of vibration-rotation bands belonging to CO_2. Thus it is not entirely certain that all of the observed discrepancies are the result of faulty description of line widths.

In conclusion we note again the very successful semi-quantitative representation which is obtained for the emissivity, even at relatively low pressures, by using the very simple method described in I. Relevant limiting emissivities are identified by the symbol E_L and are also shown in Figs. 14 to 17. The result that E_L does not differ greatly from the calculated values of E for the smallest optical densities follows from the fact that the approximate treatment given in I holds accurately for very small values of the optical density as well as at elevated total pressures. In view of the previously discussed uncertainties concerning line widths, it is recommended that emissivity estimates for practical purposes be performed by the simple method given in I, at least until it is possible to utilize a sound description of line shape and half width.

Reference to the calculated emissivity values for HCl shows again that E_L and E differ relatively less for the smallest and largest values of the optical density. In this case the treatment based on the assumption that no overlapping occurs between rotational lines would be expected to yield more reasonable estimates for the emissivity than the treatment described in I. The preceding statement follows from the fact that the spacing between the rotational lines of HCl is very large compared to the half-width unless the total pressures are exceedingly large (*viz.*, in excess of 1000 psia).

APPENDIX I. EFFECT OF VIBRATIONAL EXCITATION ON THE COLLISION DIAMETER

The collision (Lorentz) half-width b is customarily used to estimate an "optical collision diameter" ρ. It is well known that

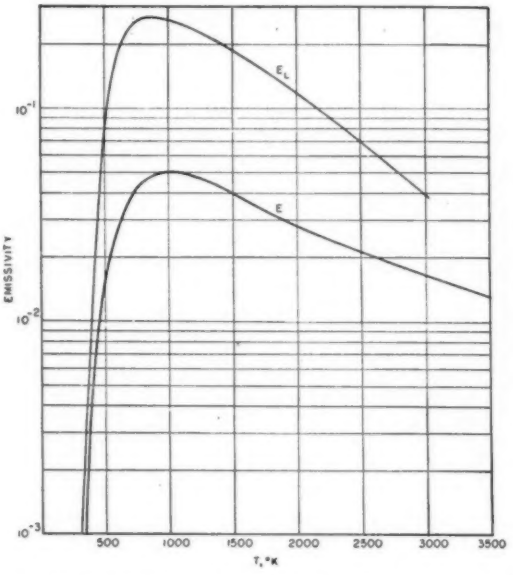


FIG. 22. Total emissivities as a function of temperature for HCl at $X=100$ cm atmos (emissivities based on treatment for non-overlapping rotational lines: E; limiting emissivities: E_L).

the square of the optical collision diameter (ρ^2) is proportional to the Lorentz half-width (b) and is related qualitatively to the square of the collision diameter of transport theory (σ_{AB}^2) for the ground vibrational energy level.[8] Unfortunately it is hardly safe to extrapolate the tenuous relation observed for molecules in their ground vibrational level to molecules in excited vibrational levels. For this reason we consider the following arguments to be really useful only for estimating the dependence of σ_{AB} on vibrational quantum number. The results concerning the probable variation of σ_{AB} with vibrational quantum number are of interest in connection with a determination of the limits of validity of calculated transport coefficients as the temperature is raised.[17] It should be noted, however, that application of the present considerations to the theoretical calculation of transport properties cannot be made

[17] S. Chapman and T. G. Cowling, *Mathematical Theory of Non-Uniform Gases* (Cambridge University Press, Cambridge, England, 1939); C. F. Curtiss and J. O. Hirschfelder, J. Chem. Phys. **17**, 550 (1949); Hirschfelder, Bird, and Spotz, J. Chem. Phys. **16**, 968 (1948).

without an analysis of the dependence of intermolecular force fields on vibrational excitation.

In order to obtain a semiquantitative estimate for the effect of vibrational excitation on σ_{AB}, we assume the simple physical model shown in Fig. 23. Although σ_{AB} is not accurately equal to $\bar{r}+2r_0+r_A+r_B$ because of distortion of the electronic charge distribution of atoms in molecules and because molecules are not spherically symmetrical, the model is probably good enough for the problem under discussion. We proceed by showing first that $2r_0$ is small and varies very slowly with vibrational excitation. Next it is shown that $\bar{r}$ also varies very slowly with vibrational quantum number.

According to the Born-Oppenheimer approximation,[18] the complete wave equation for diatomic molecules may be separated into two equations which describe the electronic and nuclear motion

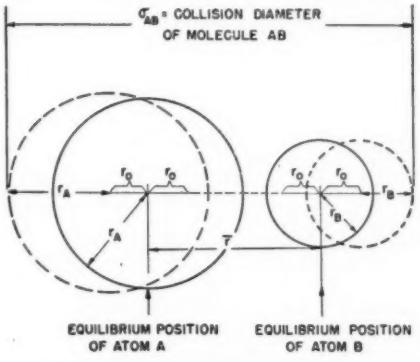


FIG. 23. Schematic representation of diatomic molecule ($\bar{r}$= equilibrium internuclear distance, obtained from spectroscopic measurements, etc.; $2r_0$=amplitude of oscillation of nuclei about their respective equilibrium positions, estimated by using the harmonic oscillator approximation; $2r_A$ and $2r_B$ represent, approximately, the equivalent collision diameters of atoms A and B).

independently. The radial part of the nuclear wave equation is found to be identical with the wave equation for the one-dimensional harmonic oscillator if we assume a Hooke's law internuclear potential function.†† In accord with the usual probability interpretation of the square of the wave function, we may equate $r_0=r_0(n)$ for the nth vibrational energy level to the distance from the equilibrium position at which the square of the radial wave function, $|\psi_n(r)|^2$, has fallen substantially to zero.

Reference to the plots of $\psi_n(r)$ or $|\psi_n(r)|^2$ *vs* r given by Pauling and Wilson[19] for $n=0$, 1, 2, 3, 4, 5, 6, and 10 shows that $r_0(n)/r_0(0)\sim 2$ for $0<n\leqslant 10$ with $r_0(0)\simeq(h/\pi^2 uc\omega_e)^{\frac{1}{2}}$. Representative values of $2r_0(0)$ for diatomic molecules are summarized in Table II.

TABLE II. Numerical values for σ_{AB} in the ground vibrational state, $2r_0$, α_e/B_e, and $2r_0/\sigma_{AB}$ for representative diatomic molecules.

Molecule	σ_{AB} (from viscosity data, in A units)[a]	$2r_0=2(h/\pi^2 uc\omega_e)^{\frac{1}{2}}$ (in A units)	α_e/B_e	$2r_0/\sigma_{AB}$
H_2	2.968	0.4942	4.922×10^{-2}	0.1665
N_2	3.681	0.1806	9.303×10^{-3}	0.0491
NO	3.470	0.1948	1.044×10^{-2}	0.0561
N_2	3.433	0.2066	1.092×10^{-2}	0.0602
CO	3.590	0.1905	9.084×10^{-3}	0.0531
Cl_2	4.115	0.2336	6.973×10^{-3}	0.0568
Br_2	3.859	0.2043	3.399×10^{-3}	0.0529
I_2	4.982	0.1989	3.133×10^{-3}	0.0399
HCl	3.305	0.4290	2.851×10^{-2}	0.1298

[a] From the values given by J. O. Hirschfelder et al., reference 17.

The dependence of $\bar{r}$ on vibrational excitation is obtained conveniently from the well-known relation[20]

$$B_v\simeq B_e-\alpha_e(n+\tfrac{1}{2}), \qquad \text{(A-1)}$$

where $B_v=h/8\pi^2 uc\bar{r}^2$ and $B_e=h/8\pi^2 ucr_e^2$. Since $\alpha_e/B_e\ll 1$, we obtain from Eq. (A-1) the result

$$\bar{r}/r_e=[1-(\alpha_e/B_e)(n+\tfrac{1}{2})]^{-\frac{1}{2}}\simeq 1+\tfrac{1}{2}(\alpha_e/B_e)(n+\tfrac{1}{2}). \qquad \text{(A-2)}$$

From the numerical values given for α_e/B_e in Table A-1 it is evident that $\bar{r}/r_e$ is only very slightly greater than unity for the values of n which are encountered in practical applications. Furthermore, reference to Table A-I shows that $2r_0(0)/\sigma_{AB}$ is small, whence it follows that the collision diameter of diatomic molecules is not a sensitive function of vibrational excitation.

[18] See, for example, L. Pauling and E. B. Wilson, Jr., *Introduction to Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York), Chapter X.

†† The refinements resulting from the use of a Morse potential are hardly warranted for the present discussion.

[19] Reference 18, pp. 74–76.

[20] See, for example, G. Herzberg, *Spectra of Diatomic Molecules* (D. Van Nostrand Company, Inc., New York, 1950), p. 106.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 8 AUGUST, 1952

Time-Temperature Dependence of Linear Viscoelastic Behavior

F. SCHWARZL AND A. J. STAVERMAN
Plastics Research Institute T.N.O., Delft, Holland

(Received February 4, 1952)

The question is treated whether in the study of linear viscoelastic behavior of a material change of temperature is completely equivalent to a shift of the logarithmic time scale or not. If it is, the material is termed "thermo-rheologically simple" (class *A*).

It is shown that, by plotting the results of a similar rheological experiment (for instance, a creep experiment) performed at different temperatures and comparing the curves obtained, one can decide whether the materials considered belongs to class *A* or not, by seeing whether the curves can be made to fit by shifting them along the axis of logarithmic time. Once the material has been decided to behave thermo-rheologically simple, one can plot the function describing the time-shift as function of the temperature.

Some examples of thermo-rheologically simple materials are quoted from the literature.

The micro-rheological conditions for thermo-rheologically simple behavior are discussed. It turns out that this behavior implies that in similar deformations at different temperatures always the same sequence of molecular events follows, whereas in materials not belonging to class *A*, not only the speed, but also the sequence of molecular processes changes when the temperature of the experiment changes. This implies, that materials of class *A* cannot from a heat treatment assume a special structure which could not be obtained by one temperature, whereas for materials of class *B* a heat treatment can indeed result in a special structure. It is reasonable to expect that thermo-rheologically simple materials will be found only among polymers containing no crystallites and no pronounced polar groups.

1. INTRODUCTION

ACCORDING to the theory[1,2] linear viscoelastic behavior of a material can be described by any one of the following functions:

$$\begin{aligned}
&\text{the relaxation function } \Psi(\ln t), && \text{(a)}\\
&\text{the creep function } \varphi(\ln t) \text{ or the creep function without flow } \Phi(\ln t)=E^{-1}\varphi(\ln t)-t/\eta, && \text{(b)}\\
&\text{the complex elasticity modulus } E^{+}(\ln\omega) \\
&\qquad =E_1(\ln\omega)+iE_2(\ln\omega), && \text{(c)} \quad (1.1)\\
&\text{the complex compliance } Z^{+}(\ln\omega) \\
&\qquad =Z_1(\ln\omega)-iZ_2(\ln\omega), && \text{(d)}\\
&\text{the relaxation spectrum } G(\ln\tau) \text{ or } L(\ln\tau) \\
&\qquad =\tau\cdot G(\ln\tau), && \text{(e)}\\
&\text{the retardation spectrum } J(\ln\tau) \text{ or } M(\ln\tau) \\
&\qquad =\tau\cdot J(\ln\tau), && \text{(f)}
\end{aligned}$$

in which E is the Hookean elasticity modulus and η the Newtonian viscosity. Of these six functions the first four can be determined immediately, namely, (a) by relaxation measurement, (b) by creep measurement, and (c) and (d) by dynamic (oscillation) measurements with prescribed deformation or force. The other two, (e) and (f), must be obtained by integral transformation, in particular (e) from (a) and (c), and (f) from (b) and (d).

As stated above, any one of these functions suffices to characterize the material completely, with no other restriction besides linearity of the behavior and constancy of temperature, the condition of linearity meaning that if all tensions are doubled, all deformations are doubled or, in other words, that Boltzmann's superposition principle is valid.

Theoretically, therefore, it is unimportant as to which function is used to characterize the material as all other functions can be calculated if one is known, and one would be inclined to use one type of measurement and characterize the material by the function following immediately from that measurement.

In practice, however, relaxation and creep experiments cannot be performed in very short times ($<10^{-3}$ sec) nor can oscillation experiments be extended easily to very long times.

Moreover, for a complete characterization of the material, it is desirable that data are obtained from a very long range of the time scale, which is not easily accessible experimentally anyway. Particularly in investigations of the relation between viscoelastic behavior and molecular structure, the influence of chemical substitutions on the behavior of the material may well be perceptible only in a range of the time scale above or below the experimental region.

As a means of extending the experimental region, so to say artificially, it is common practice to change the temperature. This procedure is based on the well-known fact that a slight increase in temperature generally increases all rate processes. Therefore, by measurements at different temperatures an extension of the time scale accessible to experiment, which exceeds everything obtainable by measurements at one temperature only, is possible. Also the temperature dependence of mechanical properties is in itself a problem of immense practical interest.

However, the interpretation of measurements at different temperatures meets with completely new difficulties. Essentially the functions (1.1) are defined only at a constant temperature. Not only is it impossible to predict with (1.1) the behavior of the material, if T changes with time, but one does not know either how to calculate the functions (1.1) at a constant temperature T_2, when they are given at a constant temperature T_1. This means that the statement that change of temperature means just a shifting of the time scale is subject to serious uncertainty.

It is the purpose of this paper to investigate what

[1] B. Gross, J. Appl. Phys. **18**, 212 (1947).
[2] B. Gross, J. Appl. Phys. **19**, 257 (1948).

conditions have to be fulfilled in order to permit calculation of the functions (1.1) at a certain temperature from measurements at another temperature.

In Sec. 2 will be given the general equations which we will use for the description of viscoelastic behavior, and in Sec. 3 the conditions which a material must obey in order to make replacement of temperature change by a shift on the time scale permissible; Sec. 4 contains the general theory of viscoelastic behavior of materials obeying the conditions of Sec. 3 (materials of class A); in Sec. 5 the difference between materials allowing temperature-time-shift (class A) and those not allowing this procedure (class B) will be discussed with respect to the molecular structure.

2. GENERAL RELATIONS

The deformation γ at constant load (i.e., the creep) of a viscoelastic material can, if we confine ourselves to the region of linear behavior, be written as a sum of three terms:

$$\gamma(\ln t, T)=\gamma_1(T)+\gamma_2(\ln t, T)+\gamma_3(\ln t, T). \quad (2.1)$$

These three terms can be discriminated phenomenologically according to their time dependence.

The first term is the ideal Hookean deformation and is independent of time:

$$\gamma_1(T)=\sigma/E \quad (2.2)$$

with σ the force and E the elasticity modulus.

The second term represents retarded elasticity and is generally strongly dependent on time and temperature:

$$\gamma_2(\ln t, T)=\sigma\cdot\Phi(\ln t, T). \quad (2.3)$$

The third term in (2.1) represents Newtonian flow

$$\gamma_3(t, T)=\sigma t/\eta. \quad (2.4)$$

It increases linearly with the time and very rapidly with increasing temperature. Attempts to account for the temperature dependence by a relation of the form

$$\eta=\text{const}\exp E_\eta/RT, \quad (2.5)$$

in which E_η should be a temperature-independent "activation energy of flow," have failed for several polymers.[3-7] For our purpose it suffices to put

$$\ln\eta=-g(T), \quad (2.6)$$

in which $g(T)$ is an arbitrary function of temperature, increasing with increasing temperature. The function of Eq. (2.5) represents a special case of $g(T)$.

From (2.4) and (2.6) we see that

$$\ln\gamma_3=\ln\sigma+\ln t+g(T) \quad (2.7)$$

or if we put

$$\ln t+g(T)=y \quad (2.8)$$

$$\gamma_3=\sigma\cdot\exp y, \quad (2.9)$$

(2.9) is not only a way of abbreviating (2.7). It also means that time and temperature appear in γ_3 through y only, which means only in the combination $\ln t+g(T)$. In other words, any change of temperature is equivalent to a shift of $\ln t$. It is this result which will be generalized so as to include γ_2 in the next section.

3. DISCRIMINATION BETWEEN MATERIALS ALLOWING TIME-TEMPERATURE-SHIFT (CLASS A) AND THOSE WHICH DO NOT (CLASS B)

We will proceed now to formulate the conditions which a material must fulfill in order to justify the procedure of replacing any change of temperature by a shift on the logarithmic time-scale. Materials fulfilling these conditions will be defined as materials of class A; they could be termed "thermo-rheologically simple" in contrast to those materials which do not allow the replacement of temperature change by a shift of the logarithmic time-scale, which we define as materials of class B.

It follows then from Sec. 1 that materials of class A are much better suited to a complete characterization over a large range of the time scale than those of class B. Moreover, the rheological behavior of materials of class A can be described for the entire measured temperature range with one function of temperature only.

Let us first consider γ_2, defined in Sec. 2,

$$\gamma_2=\sigma\Phi(\ln t, T). \quad (2.3)$$

The fact, that for materials of class A, change of temperature is equivalent to a shift on the logarithmic time-scale means that Φ depends on t and T only in the combination

$$x=\ln t+f(T), \quad (3.1)$$

in which $f(T)$ is an arbitrary function of the temperature, which must be determined experimentally. As a condition which a material must fulfill in order to belong to class A, we can state then: A function $f(T)$ must exist of such a nature, that

$$\Phi(\ln t, T)=\Phi(x). \quad (3.2)$$

In order to investigate whether or not a given material belongs to class A and at the same time to determine $f(T)$ if the material does belong to class A, the following procedure can be adopted.

One makes creep experiments at different temperatures, the temperatures being constant in each experiment. The results are plotted as Φ against $\ln t$. First one investigates whether or not the creep curves, obtained at different temperatures, can be made to coincide by shifting them parallel to the $\ln t$-axis in such a way that the points at one definite Φ-value fall together. Let the

[3] J. D. Ferry, Physics **6**, 356 (1935).
[4] J. D. Ferry, J. Am. Chem. Soc. **64**, 1330 (1942).
[5] T. G. Fox and P. J. Flory, J. Am. Chem. Soc. **70**, 2384 (1948).
[6] T. G. Fox and P. J. Flory, J. Appl. Phys. **21**, 581 (1950).
[7] H. Leaderman, *Physics of High Polymers* (Utrecht, 1951). University Press.

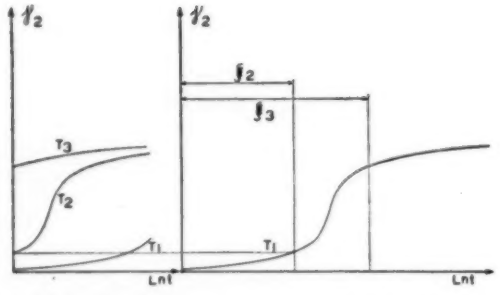


FIG. 1. Illustrating the time-temperature-shift of the creep curve of materials belonging to class A.

shifts along the $\ln t$ axis necessary to achieve this be designated by 0, ξ_2, and ξ_3 (Fig. 1) if the temperatures of measurement are T_1, T_2, and T_3. If the curves do coincide to within experimental error the first condition for belonging to class A is fulfilled. The function $f(T)$ can then be determined by plotting 0, ξ_2, and ξ_3 against T_1, T_2, and T_3 (Fig. 2).

From (3.1) and (3.2) it follows that

$$\Phi(\ln t, T)=\Phi(\ln t', T') \text{ if } \ln t+f(T)=\ln t'+f(T'). \tag{3.3}$$

So by this procedure only the variation of $f(T)$ with T can be obtained. In practice it is the variation of $f(T)$ that matters. It is convenient to extrapolate $f(T)$ towards $T=0$ and put by definition

$$f(0)=-\infty, \tag{3.4}$$

which means that at absolute zero it takes infinite time to get a finite strain.

With the introduction of (3.4) a unique value of $f(T)$ can be obtained, provided the extrapolation towards $T=0$ can be made sufficiently accurate. In practice the extrapolation comes down to looking for a negative power of T—negative in order to satisfy (3.4)—which varies like the experimental $f(T)$ in the experimental region of T. If viscoelastic behavior is considered as a rate phenomenon in Eyring's sense with a constant activation energy,

$$f(T)=-a/T. \tag{3.5}$$

It has been found, however, that better descriptions are given by*

$$f(T)=-a/T^2 \tag{3.6}$$

for polyisobutylene[3–7] and

$$f(T)=-(a/T-T_\infty) \tag{3.7}$$

for rubber[8] and polymethylmethacrylate.[9] The function $f(T)$ of Eq. (3.7) does not satisfy (3.4), but instead $f(T_\infty)=-\infty$. It should be remarked, however, that functions such as (3.6) and (3.7) are only used to describe the measurements in a certain limited region of temperature. There exists no real justification for extrapolation outside the region used in the measurements.

Now the condition outlined in this section is necessary, but not sufficient, to place a material in class A, as the definition of such material implies that for the entire rheological behavior, as described by any of the functions (1.1), a change of temperature is equivalent to a shift of $\ln t$. We will see in the next section that this definition is satisfied only if the temperature dependence of γ_3 is identical with that of γ_2 or if

$$f(T)=g(T). \tag{3.8}$$

Of course the existence of a function $f(T)$, even if different from $g(T)$, means a great simplification of the description of viscoelastic behavior, and one could propose to define a special class for materials of which γ_2 can be described with one function $f(T)$, but with

$$f(T)\neq g(T).$$

Micro-rheological considerations, however, make us feel that the division of materials as proposed here has some relation with fundamental properties of these materials.

4. GENERAL THEORY OF THERMO-RHEOLOGICAL SIMPLE MATERIALS (CLASS A)

Now consider the complete deformation of a material showing the time-temperature shift in the viscoelastic strain γ_2. Collecting the results of formula (2.1), (2.2), (2.9), and (3.2) we get for the total strain (creep curve)

$$\gamma/\sigma=\gamma_1/\sigma+\gamma_2(x)/\sigma+\gamma_3(y)/\sigma. \tag{4.1}$$

The first term on the right-hand side of (4.1) equals the reciprocal Hookean elasticity modulus E^{-1} and depends only slightly on temperature. To simplify our expressions, we take γ_1 as independent of temperature, which is permissible if compared with the large temperature dependence of the other terms. The second term possesses a time-temperature shift of the form

$$x=\ln t+f(T),$$

while the third term possesses one of the form

$$y=\ln t+g(T).$$

It is clear, therefore, that the entire viscoelastic behavior can be described by one single function of

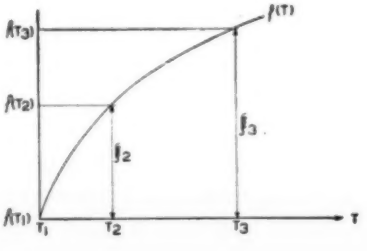


FIG. 2. The time-temperature relationship determined by plotting the time shift against the temperature differences.

* An excellent survey of the data is given by Dr. H. Leaderman in his lectures in Delft, Utrecht, and Groningen. Of these lectures there exists only a stencilled report, which contains also many unpublished experimental results.

[8] Conant, Hall, and Lyons, J. Appl. Phys. **21**, 499 (1950).

[9] Bischoff, Catsiff, and Tobolsky, ONR Technical Report RLT-2, Princeton University (December 1951).

temperature only if either

$$g(T)=f(T) \tag{4.2a}$$

or

$$\gamma_3(y)\equiv 0; \tag{4.2b}$$

that is to say, there is no Newtonian flow at all.

If one of these conditions is fulfilled together with (3.2), we speak of a thermo-rheologically simple material (class A).

We want now to investigate the properties of materials of class A. First we ask for the shape of the retardation spectrum, which is defined by

$$\begin{aligned}&\Phi(\infty,T)-\Phi(\ln t,T)\\&=\frac{1}{\sigma}\gamma_2(\ln t,T)=\int_{-\infty}^{+\infty}M(\ln\tau,T)e^{-t/\tau}d\ln\tau\\&=\int_{-\infty}^{+\infty}M(\ln\tau,T)\exp[-\exp(\ln t-\ln\tau)]d\ln\tau,\end{aligned} \tag{4.3}$$

in which the integration is understood to be performed with T constant.

Now we put $\ln\tau+f(T)=z$ and $\ln t+f(T)=x$, and replace $\ln\tau$ everywhere by $z-f(T)$ and $\ln t$ by $x-f(T)$ and, therefore, $M(\ln\tau,T)$ by $M'(z,T)$:

$$\frac{1}{\sigma}\gamma_2(x)=\int_{-\infty}^{+\infty}M'(z,T)\exp[-\exp(x-z)]dz,$$

in which $M'(z,T)$ does not depend on t or x, but may or may not depend on T explicitly, i.e., except through z. Now after integration and inserting the limits the variable z disappears and we obtain a function of x, which depends on T explicitly if M' does so. However, this function is equal to $(1/\sigma)\gamma_2(x)$, which we know does not depend on T except through x, so M' cannot depend on T either, or

$$M'(z,T)\equiv M(\ln\tau,T)=M(z), \tag{4.4}$$

which means that the retardation spectrum obeys exactly the same time-temperature shift as the creep function; the whole spectrum is shifted to shorter times in changing to higher temperatures, without changing its shape.

The next characteristic functions we are interested in are the real and imaginary parts of the dynamic compliance, which are given by the equations

$$\left.\begin{aligned}Z_1(\ln\omega,T)&=\frac{1}{E}+\int_{-\infty}^{+\infty}M(\ln\tau,T)\frac{d\ln\tau}{1+\omega^2\tau^2}\\Z_2(\ln\omega,T)&=\frac{1}{\omega\eta}+\int_{-\infty}^{+\infty}M(\ln\tau,T)\frac{\omega\tau d\ln\tau}{1+\omega^2\tau^2}\end{aligned}\right\}. \tag{4.5}$$

If we write here $f(T)-\ln\omega=x$ and $f(T)+\ln\tau=z$ and if we remember that Eqs. (2.6) and (3.8) yield $1/\omega\eta=e^x$, we arrive at the results

$$\begin{aligned}Z_1&=\int_{-\infty}^{+\infty}\frac{M(z)\,dz}{1+\exp[2(z-x)]}+\frac{1}{E}=Z_1(x)\\Z_2&=\int_{-\infty}^{+\infty}\frac{M(z)\,dz}{1+\exp[2(z-x)]}\exp[z-x]+e^x=Z_2(x).\end{aligned} \tag{4.6}$$

These equations show that the dynamic functions are only dependent on x, but not on $\ln t$ or T separately.

To find the dependence of the relaxation spectrum we use a relationship between the two spectra, which has been given by Gross and Pelzer:[10]

$$L(\ln\tau,T)=\frac{M(\ln\tau,T)}{\left[1+\dfrac{E\tau}{\eta}+\tau\displaystyle\int_0^{\infty}\frac{M(\ln u,T)\,du}{u(\tau-u)}\right]^2+[\pi M(\ln\tau,T)]^2}. \tag{4.7}$$

If we assume $M(\ln\tau,T)=M(z)$ we see at once that the numerator and three terms of the denominator only depend on z. The only term which asks for further investigation is the integral

$$\begin{aligned}\tau\int_{-\infty}^{\infty}\frac{M(\ln u,T)\,d\ln u}{\tau-u}&=\int_{-\infty}^{+\infty}\frac{M(\xi)\,d\xi}{1-u/\tau}\\&=\int_{-\infty}^{+\infty}\frac{M(\xi)\,d\xi}{1-\exp[\xi-z]},\end{aligned} \tag{4.8}$$

where we have put $\xi=\ln u+f(T)$ as the new integration variable. Equation (4.8) shows that the integral is a function of z only, and therefore the whole expression (4.7) is only dependent on z. This means that the relaxation spectrum obeys our time-temperature shift

$$L(\ln\tau,T)=L(z). \tag{4.9}$$

At last it is readily seen that the relaxation function

$$\begin{aligned}\Psi(\ln t,T)&=\int_{-\infty}^{+\infty}L(z)[1-e^{-t/\tau}]d\ln\tau\\&=\int_{-\infty}^{+\infty}L(z)[1-\exp(-\exp\{x-z\})]dz\end{aligned} \tag{4.10}$$

obeys our relationship as well as the dynamic elasticity modulus.

[10] B. Gross and H. Pelzer, J. Appl. Phys. **22**, 1035 (1951).

Collecting our results we get the following important statement: If a linear viscoelastic material belongs to class A, then all its characteristic functions fulfill the same time-temperature shift and can be represented as a function of the reduced time:

$$\begin{aligned} &x=\ln t+f(T) && x=-\ln\omega+f(T) \\ &z=\ln\tau+f(T) \end{aligned} \tag{4.11}$$

$$\begin{aligned} &\Psi(\ln t,\ T)=\Psi(x) && E_1(\ln\omega,\ T)=E_1(x) \\ &L(\ln\tau,\ T)=L(z) && E_2(\ln\omega,\ T)=E_2(x) \\ &\Phi(\ln t,\ T)=\Phi(x) && Z_1(\ln\omega,\ T)=Z_1(x) \\ &M(\ln\tau,\ T)=M(z) && Z_2(\ln\omega,\ T)=Z_2(x). \end{aligned} \tag{4.12}$$

It is interesting to mention that the spectra $G(t)$ and $J(t)$ in the original time scale do not fulfill the time-temperature shift. It seems, therefore, that the spectra in the logarithmic time scale are preferable to these.

If our material shows a nonvanishing viscous flow, a second interesting representation can be shown: because of $x=\ln(1/\omega\eta)=\ln(t/\eta)$, the curves, belonging to different temperatures, must fall together if they are plotted as functions of t/η or $\omega\eta$ instead of t or ω:

$$\begin{aligned} &\Psi(\ln t,\ T)=fct(t/\eta) && E_1(\ln\omega,\ T)=fct(\omega\eta) \\ &L(\ln t,\ T)=fct(t/\eta) && E_2(\ln\omega,\ T)=fct(\omega\eta) \\ &\Phi(\ln t,\ T)=fct(t/\eta) && Z_1(\ln\omega,\ T)=fct(\omega\eta) \\ &M(\ln t,\ T)=fct(t/\eta) && Z_2(\ln\omega,\ T)=fct(\omega\eta). \end{aligned} \tag{4.13}$$

This representation has been found experimentally for polyisobutylene.[11,12]

5. MICRO-RHEOLOGICAL DISCUSSION

The molecular cause for viscoelastic behavior is that the deformation produces in the molecular configurations changes of high free energy, which disappear partly or wholly with a measurable rate. During the viscoelastic process there is an exchange of energy from configurational free energy into heat in which the forces, aroused by the configurational free energy, act as driving forces for the rate process in which the energy is dissipated. The characteristic time for a particular molecular process can be calculated as the quotient of a characteristic elasticity force (force per unit displacement) and a characteristic viscosity (force per unit rate of displacement).

Temperature change can affect the micro-rheological events in two ways:

In the first place, increase of temperature will increase the velocity of all molecules and atoms, and therefore of all rates of configurational change.

In the second place, the changes of free energy with a given change of molecular configuration will vary with the temperature and, therefore, the driving force will vary with temperature.

The latter effect, which is by far the smaller, can lead to decrease of the forces with increasing temperature, when the intermolecular forces are the main source of the configurational free energy, but it leads to increase of the force with increase of temperature, when the molecular kinetic energy is the main source of the configurational free energy as is the case with such highly elastic substances as rubber.

The first effect, which exceeds the second many times, can be considered as the effect of the temperature on the "viscosities" of the molecular processes, and it is well known that viscosities are generally strongly dependent on temperature.

We see now what is needed in order that a material behaves thermo-rheologically simple (class A). If we neglect the effect of the temperature on the forces we see that all characteristic times will be increased in the same proportion if all the viscosities show the same dependence on temperature. Indeed, if we designate one particular molecular process by the index i, then, if

$$\eta_i(T_1)/\eta_i(T_2)=f(T_1T_2) \text{ independent of } i \tag{5.1}$$

(η_i being the "characteristic viscosity" of process i), all molecular processes will be accelerated in the same proportion at a certain change of temperature. Condition (5.1) is the micro-rheological counterpart of (3.2) with (3.1) and (3.8), by which materials of class A are defined.

However, (5.1) means that the apparent activation energy of all molecular rate-processes must be identical. It is clear that this can be realized only if all molecular interaction forces are of the same kind, as is the case in hydrocarbons, such as polyisobutylene and rubber.

From the chemical point of view we may expect thermo-rheological simple behavior only in case of non-crystalline polymers without strongly polar groups, since both the changes of crystalline structure and of interaction of polar groups will certainly depend on the temperature in a way different from that of the viscosities of flow in the amorphous, nonpolar parts of the material.

From this discussion it is also clear that temperature change represents a shift of $\ln t$ and not of t itself or another function of t, since all rates, and therefore all times, are affected in the same proportion to their absolute values by change of temperature.

We are now in a position to understand why materials of class A are not only more conveniently amenable to a thermo-rheological description, but are essentially less complicated than those of class B.

If we consider the molecular processes, which all together constitute the macroscopic rheological behavior as formally independent, it is clear that we must reckon with a more or less strong interaction between these processes. This means that the rate of a certain process i will depend on whether or not a particular event E_j

[11] R. S. Marvin, Interim Report Natl. Bur. Standards, April, 1951.

[12] R. S. Marvin, Fitzgerald, and Ferry, J. Appl. Phys. **21**, 197 (1950).

belonging to the course of events called process j, has occurred. After every event belonging to process j, the process i finds, so to say, a new material.

Now, if the rates of all processes are changed in the same proportion, the time needed for a certain macroscopic process may change very much, but *the sequence of micro-rheological events remains unaltered*. This means that in materials of class A at different temperatures essentially the same molecular processes take place only with different speed. On the contrary in materials, belonging to class B, the sequence of molecular events is different at different temperatures. This, however, means that at a certain temperature T_1 certain molecular events may occur that do not occur at all at a temperature T_2, since as a result of changes in the sequence of events a certain molecular process may at a certain instant in its course find at T_1 a molecular configuration completely different from that at T_2 in the same instant, so it may be that the process takes at T_1 a course completely different from that at T_2.

The distinction between materials of class A and B is not only of great theoretical, but also of practical interest. Indeed if a material of class A is subjected to a certain heat treatment under tension or after deformation, the changes of temperature cannot give rise to new molecular structures, since the rates of all molecular processes are affected similarly. On the other hand, in materials of class B certain molecular configurations with desired properties may be obtained by a heat treatment, which cannot be obtained at constant temperature or with another heat treatment.

In concluding we will indicate a possible generalization of our treatment in connection with a micro-rheological model proposed by Alfrey.[13] Alfrey proposes a spectrum of the type

$$M=\begin{cases} 0 & \tau<\tau_1 \\ c\ln\tau/\tau_1 & \tau_1<\tau<\tau_2 \\ 0 & \tau_2<\tau \end{cases} \quad \text{with } c \text{ proportional to } T^{-1},$$

in which not only all viscosities show the same temperature dependence (activation energy E independent of chain length) but in which also the elasticities depend on temperature by a factor T (rubber elasticity). As the characteristic times depend on the quotient of viscosity and elasticity, all characteristic times are changed in the same proportion by temperature if all viscosities and all elasticities have the same temperature dependence and, therefore, in Alfrey's model the sequence of molecular events is not affected by changes of temperature. But owing to the fact that the elasticities change with temperature, the shape of the spectrum depends on the temperature; in Alfrey's case the height of the spectrum M is proportional to T^{-1}. This means that macroscopically Alfrey's model does not belong to class A though microscopically it does. In fact, in Alfrey's model, as a result of c being proportional to T^{-1}, the entire creep curve is proportional to T^{-1}. It is not very difficult to show that therefore we get, instead of (3.2),

$$T[\Phi(\infty)-\Phi(t, T)]=F(x). \tag{5.2}$$

Now, of course, it is not essential that the influence of temperature on the elasticities should be given by the factor T, as in fact for any other function of T the same treatment should hold.

In order to save the micro-rheological sequence of events it is not even necessary that all the viscosities and all the elasticities separately obey the same temperature dependence. It should be sufficient if (5.1) were replaced by the more general expression

$$[\tau_i(T_1)/\tau_i(T_2)]=[\eta_i(T_1)G_i(T_2)/\eta_i(T_2)G_i(T_1)] \quad \text{independent of } i.$$

However, in all cases where the spectra are not only shifted, but also change their shape on change of temperature, it will be very difficult to recognize macroscopically whether the sequence of events is preserved. Besides, the influence of temperature on viscosities exceeds that on elasticities to such an extent that the latter will generally not be noticed at all in the very insensitive creep curve. Moreover, even if the spectrum is only shifted and not changed this might be due to a fortuitous compensation of effects rather than to the preservation of molecular processes.

Therefore we think it most appropriate for the moment to define rheological simple substances by the shift—without change—of the macroscopic characteristic functions, although we realize that this definition does perhaps not comprise all materials with invariable sequences of molecular events and even could eventually include a few outside that class.

The authors are indebted to Dr. H. Leaderman (Washington, D. C.) for information and stimulating discussions.

[13] T. Alfrey, J. Chem. Phys. **12**, 374 (1944). T. Alfrey and P. M. Doty, J. Appl. Phys. **16**, 700 (1945).

An Estimate of the Energies of the Positive Ions in a Negative Point-to-Plane Corona

G. L. WEISSLER AND MARK SCHINDLER
Physics Department, University of Southern California, Los Angeles, California
(Received February 18, 1952)

An attempt has been made to measure the energies with which positive ions arrive at the cathode of a negative point-to-plane corona in N_2. A hairpin thermocouple was used as the point electrode. It was heated by the corona current, and its thermal emf was recorded by a galvanometer. After appropriate calibration, the potential across the Crookes darkspace was determined to be about 300 volts, and the average energy with which ions arrive at the corona point to be about 15 ev.

INTRODUCTION

IN order to further the understanding of negative point-to-plane corona, the basic mechanisms contributing to its propagation have been analyzed in detail by Loeb.[1] For the sake of simplified presentation, these mechanisms may be divided into two categories: (1) those occurring in the volume of gas, and (2) those active at the cathode surface. The latter involve the liberation of secondary electrons (a) by positive ion bombardment, (b) by photons, and (c) by metastables. These processes are very sensitive to the character of the cathode surface which in itself may change during the course of a discharge. Loeb[2] has also made a comprehensive analysis of "The Role of the Cathode in Discharge Instability" and discusses in particular the manner in which the stream of positive ions flowing into the negative point of a corona discharge may alter its surface significantly. Such changes are indicated by the measurements of work functions of gas-coated surfaces.[3] This in turn must lead to correspondingly changed yields of the above-mentioned cathode mechanisms.

For this reason it was thought important to undertake an experimental study of the energies with which positive ions will arrive at the point-electrode of a negative point-to-plane corona. This was attempted in the following manner. Instead of using a point-electrode made of a single wire, an iron-constantan thermocouple was employed as shown in Fig. 1. Its junction was right at the tip of the narrow hair-pin and was polished carefully to resemble as closely as possible a hemispherically capped point of about 0.5-mm diameter. With a gap length of about 3.5 cm the geometry was sufficiently close to previous corona research[4] to give rise to the same current-voltage characteristics in nitrogen. This is of importance since it has been shown[4] that small amounts of negative-ion-forming impurities such as oxygen will radically alter the characteristic curves, owing to the appearance of Trichel pulses that limit the discharge current periodically by negative ion space charge formation.[1] Therefore, in the experiments described here, a negative point corona in nitrogen was used because the thousandfold greater currents[4] (compared to air) were more likely to produce a measurable heating effect on the corona thermocouple point (hereafter referred to as c.th.c.). If the c.th.c. is calibrated separately in terms of heating power supplied *versus* deflection of the c.th.c. galvanometer, then an estimate of the positive ion energies can be made. This in turn will make it possible to obtain the potential difference between the negative glow and the cathode (Crookes dark space) which has been estimated[1,2] to be about 500 volts. Furthermore, such data would determine a more definite value of γ_i, the second Townsend coefficient, measuring the number of electrons released from the cathode per incident positive ion.

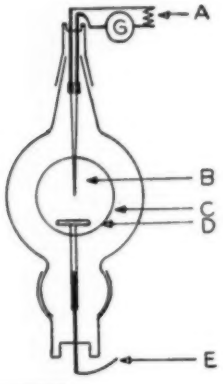


FIG. 1. Corona chamber. *A*: galvanometer circuit; *B*: corona thermocouple (c.th.c.) point; *C*: flat window; *D*: plane electrode; *E*: high voltage connection.

APPARATUS AND EXPERIMENTAL PROCEDURE

The corona gap was enclosed in a Pyrex tube, Fig. 1, provided with a plane window for visual observations and connected to a conventional vacuum system and gas purification train. Procedures for cleaning the tube and filling it with pure N_2 have been described elsewhere.[4] The c.th.c. was made of iron and constantan using B.S. wire gauge No. 25 corresponding to a wire diameter of 0.45 mm. An attempt was made to shape the junction into a hemispherical cap of 0.5-mm diameter. The characteristic curves of voltage *versus* current were obtained for coronas at three different nitrogen pressures: $p_1=703$ mm Hg; $p_2=301$ mm Hg; and $p_3=149$ mm Hg. They indicated by comparison with previous results[4] that Trichel pulses were absent and the gas was sufficiently pure. Measurements of the geometry of the

[1] L. B. Loeb, J. Appl. Phys. **19**, 882 (1948).
[2] L. B. Loeb, Phys. Rev. **76**, 255 (1949).
[3] G. L. Weissler and R. W. Kotter, Phys. Rev. **73**, 538(A) (1948); G. L. Weissler and T. N. Wilson, Phys. Rev. **76**, 591(A) (1949).
[4] G. L. Weissler, Phys. Rev. **63**, 96 (1943).

visual corona were taken with the aid of a calibrated telemicroscope. The heating of the c.th.c. due to the discharge currents was registered by a galvanometer of sensitivity 0.01 μa/cm. The c.th.c. was held near ground, whereas the plate electrode was at a high positive potential. Corrections were made for the fact that the discharge current had to pass through either one leg or the other of the c.th.c., and measurements were then taken of the deflections of the c.th.c. galvanometer at various corona currents and three nitrogen pressures as presented in Fig. 2.

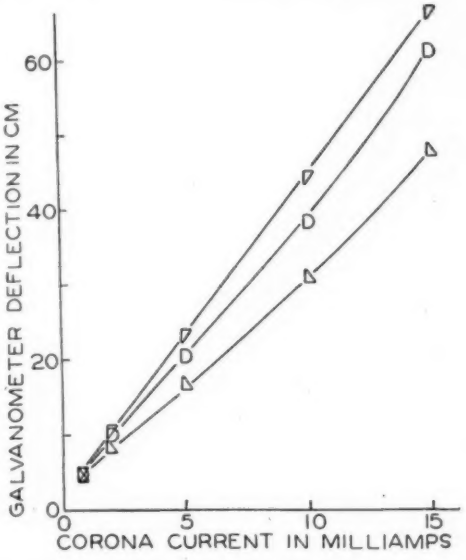


FIG. 2. Corona thermocouple galvanometer deflections *versus* corona currents. Upper curve: N_2-pressure $p=149$ mm Hg; middle curve: $p=301$ mm Hg; lower curve: $p=703$ mm Hg.

CALIBRATION PROCEDURE

In order to determine the energy reaching the c.th.c. during the discharge, it was necessary to find what meaning the galvanometer deflections had with respect to the power input to the point. In a separate apparatus (Fig. 3) different but known amounts of heat could be conducted to the same c.th.c. point, and the resultant deflections could be observed. A "grain-of-wheat" lamp filament without its glass envelope was mounted next to a fine wire reference thermocouple (hereafter referred to as r.th.c.), both were cemented together with electrically insulating cement, and the c.th.c. could be held in contact with this assembly or separated from it. In order to allow for different areas of contact of the c.th.c. point with the heater element corresponding to the observed corona dimension, a small amount of amalgam, which could easily be shaped to the correct size, was placed on top of the cement. The supports of the heater assembly were made as small as possible in order to minimize heat losses as well as shorten the time necessary to reach equilibrium.

In the actual measurements arbitrary zeroes were established for the c.th.c. as well as the r.th.c. galvanometers before heating the filament. The tip of the c.th.c. was then moved to within 0.2 mm of the amalgam, and a current was passed through the heater element corresponding to a power W_1. This current was of such a value that the r.th.c. would produce a galvanometer deflection in its sensitive range. In addition, this heating would deflect the c.th.c. galvanometer by d_1 cm.

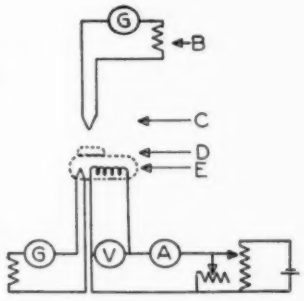


FIG. 3. Calibration equipment. *A* and *V*: current meter and volt meter to measure power to heater *E*; adjacent to *E* is the reference thermocouple (r.th.c.) with its galvanometer *G*; *D*: amalgam to provide contact between *E* and *C*, the corona thermocouple (c.th.c.) point; *B*: c.th.c. galvanometer circuit.

The next step was to establish contact between the c.th.c. and the heater by way of the amalgam. This additional path for heat loss caused a cooling and a decreased deflection of the r.th.c. galvanometer. The power to the filament was increased to a value W_2 until the r.th.c. indicated the same temperature as before contact and the changed c.th.c. galvanometer reading d_2 was noted.

In principle, if no disturbing effects were present, these measurements would lead directly to a computation of the c.th.c. galvanometer deflections as a function of power to the tip of the thermocouple. This relation would in turn enable one to establish how much energy the corona discharge transferred to the point electrode and finally to compute from this the energies of the positive ions reaching the point. However, owing to the complexities of the laws relating to heat loss, a more detailed analysis of the data was needed.

The heat losses from the point of the c.th.c. are due to (a) conduction through the junction and along the leads of the thermocouple; (b) radiation from the exposed surface; and (c) convection from the wire surfaces. These three processes depend upon different powers of

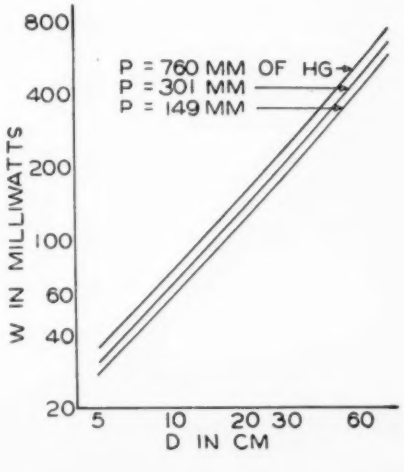


FIG. 4. Calibration graph. Power *W* to the corona thermocouple (c.th.c.) point *versus* c.th.c. galvanometer deflection *d* for three N_2 pressures.

the temperature and are represented as follows: $W_c=C_c(T-T_0)$ for conduction; $W_r=C_r(T^4-T_0^4)$ for radiation; and $W_v=C_v(T-T_0)^{1.25}(P/760)^{\frac{1}{2}}$ for convection.[5,6] Here the W's are the power losses in watts, P the pressure in mm of mercury, the C's are constants to be determined, and T, T_0 are the highest temperature at the point of the c.th.c. and room temperature, measured on the absolute scale. Since the thermocouple response in the region covered by the corona data is quite linear with respect to temperature, D and D_0 may be substituted for T and T_0 where the D's are the c.th.c. galvanometer deflections measured on an absolute temperature scale. In a more detailed analysis[7] the Fourier equation of heat flow was solved for this particular case, and the arbitrary constants were determined by curve fitting. After all corrections had been made, a calibration curve was obtained (Fig. 4) representing the power W in milliwatts *versus* the c.th.c. galvanometer deflections d for the range covered in the corona measurements. By using Figs. 2 and 4 it was then possible to obtain the absolute energy reaching the c.th.c.

TABLE I. Total voltage drop across Crookes dark space *vs* pressure and current.

P in mm Hg	Corona current in ma 0.08	0.2	0.5	1.0	1.5
703	390	300	250	250	270
301	370	310	280	280	350
149	350	300	290	300	320

EXPERIMENTAL RESULTS AND DISCUSSION

The energy exchange per unit time at the point resulting from the corona current is due to (a) nitrogen ions striking the point, W_+; (b) neutral N_2 molecules striking the point with energies exceeding kT as a result of collisions with positive ions, W_0; (c) electrons leaving the point, W_-; and (d) metallic particles sputtered from the point as well as photons and metastable N_2 molecules reaching the point, W_s. The sum of these energies

$$W=W_++W_0-W_-\pm W_s,$$

which has been found to be positive, is dissipated by convection and radiation from the c.th.c. as discussed previously.

The measured corona current reaching the point $I=I_++I_-$ is due to positive ions I_+ and electrons I_-.

[5] E. B. Steinberg, Elec. Mfg. **46**, (1948).
[6] L. M. K. Boelter, *et al.*, *Heat Transfer Notes* (University of California Press 1948), pp. XII-13, II-7.
[7] Mark Schindler, Master's thesis, The University of Southern California Library, June, 1952.

TABLE II. Positive ion energies (in units of ev) *vs* pressure and current.

P in mm Hg	Corona current in ma 0.08	0.2	0.5	1.0	1.5
703	19	15	13	13	14
301	18	15	14	14	18
149	20	15	15	15	16

This latter current, I_-, may be assumed negligibly small because of the small value of the second Townsend coefficient, $\gamma<10^{-2}$. The energy of the positive ions is gained mainly while passing through the Crookes dark space and is partly dissipated in collisions with neutral molecules. However, if the negative glow is wide compared to the thickness of the Crookes dark space, as is usual, conservation of momentum indicates that most of this energy would still be transferred to the point. The contribution of the slow sputtering as well as that of the photons and metastables may be neglected in so far as these energy considerations are concerned. The cooling of the point due to the energy carried away by electrons has also been neglected. In substance, the heating of the point electrode is then due to its bombardment by positive ions and neutral gas molecules. If one makes the initial assumption that the heating is due to positive ions only, then the heat developed is $W=I_+V$, where I_+ is very closely equal to the total measured corona current and V the potential through which the ions have fallen. If W is divided by I_+, this potential V which corresponds to the voltage across the Crookes dark space is obtained and given in Table I.

In actuality the positive ions do not gain an energy equal to this potential difference V because of collisions with neutral molecules. The Crookes dark space is inversely proportional to the gas pressure P and has been measured with a telemicroscope. When reduced to atmospheric pressure values close to Loeb's figure[1] of 2×10^{-4} cm were obtained. This corresponds to about 20 molecular mean free paths, or 20 collision in which the positive ions will transfer their energies to neutral molecules. If the potential difference V is divided by this number of collisions, then an estimate of the energies with which positive ions arrive at the point can be made as shown in Table II.

The probable accuracy of these measurements was largely determined by an estimate of the errors involved in the energy calibration of the c.th.c. and in the determination of the thickness of the Crookes dark space. A rough guess of the accuracy of the data in Table I is 50 percent.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 8 AUGUST, 1952

On the Theory of Antenna Beam Shaping*

ALLEN S. DUNBAR
Stanford Research Institute, Stanford, California
(Received February 18, 1952)

The diffraction pattern of an aperture with amplitude distribution $A(x)$ and phase function $\psi(x)$ is examined for two cases: (1) controlled variation of the amplitude distribution with known uniform phase; (2) controlled variation of phase with known amplitude distribution. For Case 1, it is shown that a symmetrically tapered amplitude distribution produces a diffraction pattern with reduced secondary lobes, while an asymmetric amplitude distribution produces a shaped pattern. Thus, antenna beam shaping may be achieved by controlled variation of amplitude distribution in an aperture in which there is uniform phase. For Case 2, it is shown that a specified beam shape may be approximated by controlled variation of phase in the aperture when the amplitude distribution is known. The method of calculation is an extension of the method originally formulated by Chu for calculating cylindrical reflectors to transform a given primary radiation pattern to a specified shaped pattern. A general formulation of Chu's method for an amplitude distribution on a curved surface is developed. The theoretical results are applied to the design of progressive-phase antennas and experimental results given.

I. INTRODUCTION

SHAPED-BEAM antennas, their theory, and the synthesis of specified beam shapes have been problems of considerable interest for some years. Early in World War II, when the need for the cosecant-squared beam for submarine search by airborne radar was urgently felt, modified reflector design, extended feeds in parabolic reflectors, and linear arrays to produce shaped-beams were studied experimentally at the Radiation Laboratory of Massachusetts Institute of Technology, and Telecommunications Research Establishment in Britain.[1] At about the same time, Levinson[2] at M.I.T. and Woodward[3] at T.R.E. attacked the problem theoretically and obtained Fourier-series solutions for the amplitude and phase of the electric field in an aperture to radiate a shaped beam.

Using an entirely different approach, Chu[4] devised a method for calculation of shaped-reflector design to transform a given primary radiation pattern to the desired shaped beam. This method is based on the principle of conservation of energy and geometrical optics. The principles of the procedure of calculation were applied by others at M.I.T., T.R.E., and elsewhere, to lenses and doubly curved reflectors.[5]

The theory of the linear array and the calculation of feeding coefficients for the synthesis of shaped beams was studied by Schelkunoff[6] at Bell Telephone Laboratories, Eaton, Riblet, and Eyges[7] at M.I.T., and MacFarlane[7] at T.R.E. More recently, Taylor and Whinnery[8] have treated the problem from the point of view of potential theory.

It is the purpose of this paper to express a theory of the formation of shaped beams and to apply this theory to progressive-phase antennas. An extension of Chu's method of calculation of shaped-beam antennas to progressive-phase antennas will be described, and a general formulation of the method will be developed in terms of an aperture and an assumed aperture distribution.

II. DEFINITION OF APERTURE AND APERTURE DISTRIBUTION

In describing the properties of an antenna, it is convenient to speak of the aperture or effective aperture of that antenna. In order to arrive at a definition of such an aperture, let us erect a mathematical plane of infinite extent which separates a region of space containing a source of electromagnetic energy from the region on the other side of the plane into which the source radiates the electromagnetic energy. If the power flow over all but some finite area of the plane is substantially zero, then that part of the plane through which the power flow is restricted is called the aperture. This definition applies equally well to a plane close to and in front of a paraboloidal or lens antenna or to an array of dipoles or slots in a transmission line.

The aperture distribution is the electric field or amplitude of the electromagnetic-wave disturbance in the aperture with regard to polarization and phase. With reference to the rectangular-coordinate system of Fig. 1, let the x axis be the aperture plane. The region above the x axis is free space (positive y direction). For simplicity we shall consider only fields that have no z dependence. Let a plane wave, restricted by the

* This work was supported in part by Air Force Contract No. AF 19(122)78.

[1] S. Silver, *Microwave Antenna Theory and Design* (McGraw-Hill Book Company, Inc., New York, 1944), chapter 13, Radiation Laboratory Series, Vol. 12; D. W. Fry, J. Inst. Elec. Engrs. (London) **93**, Part III-A, Sec. 10, 1505–1509 (1949).

[2] R. C. Spencer, Radiation Laboratory, M.I.T., Report 54-24 (1943).

[3] P. M. Woodward, J. Inst. Elec. Engrs. (London) **93**, Part III-A, Sec. 10, 1554–1558 (1946).

[4] L. J. Chu, Research Laboratory of Electronics, M.I.T., Technical Report 40 (1947).

[5] S. Silver, see reference 1; D. W. Fry, see reference 1; L. J. Chu, see reference 4; A. S. Dunbar, Proc. Inst. Radio Engrs. **36**, 1289–1296 (1948); D. W. Fry and F. K. Goward, *Aerials for Centimetric Wavelengths* (Cambridge University Press, Cambridge, 1950).

[6] S. A. Schelkunoff, Bell System Tech. J. **22**, 80–107 (1943).

[7] S. Silver, see reference 1, chapter 9.

[8] T. T. Taylor and J. R. Whinnery, J. Appl. Phys. **22**, 19–29 (1951).

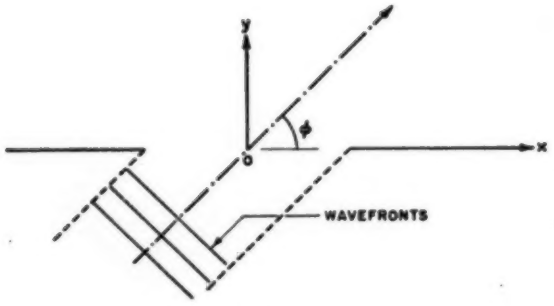


FIG. 1.

dashed lines and indicated by the wave fronts, be incident on the plane at the angle φ. The phase of the disturbance in the aperture is $(2\pi x/\lambda)\cos\varphi$ and, hence, the disturbance may be considered as a traveling wave in the aperture with phase velocity

$$v=c/\cos\varphi, \tag{1}$$

where c is the velocity of the electromagnetic wave in free space. In the case of a surface wave, of course, $v<c$, and the angle φ must be regarded as imaginary.

Consider now a wave restricted as shown in Fig. 2, but with nonuniform phase fronts, incident on the aperture. A narrow region of the wave, of width dx in the aperture, is incident at the angle φ; but, in general, the angle of incidence varies with the coordinate x in the aperture, and it is written $\varphi(x)$. Let us define a function, $\psi(x)$, such that $(2\pi/\lambda)\psi(x)$ is the phase at any point x in the aperture with reference to the origin. Thus,

$$\psi(x)=\int_{-a/2}^{x}\cos\varphi(x)dx. \tag{2}$$

Obviously, if $\varphi(x)$ is constant, as is the case in Fig. 1, the phase reduces to

$$(2\pi/\lambda)\psi(x)=(2\pi x/\lambda)\cos\varphi.$$

III. THE DIFFRACTION INTEGRAL AND DIFFRACTION PATTERNS

The far-field or Fraunhofer diffraction pattern of the aperture may be expressed in terms of the amplitude distribution $A(x)$ and the phase function $\psi(x)$ defined by Eq. (2). Thus, the amplitude of the wave disturbance at a point P in the far field in the direction θ from the origin, as shown in Fig. 3, is given by

$$E(\theta)=f(\theta)\int_{-a/2}^{a/2}A(x)\exp\{ik[x\cos\theta-\psi(x)]\}dx, \tag{3}$$

in which $k=2\pi/\lambda$ and $f(\theta)$ is the form factor that depends upon the polarization of the wave in the aperture.

In order to examine the shape of the diffraction pattern we will consider two cases: (1) The phase function in the aperture is known to be uniform and we wish to determine the effect of controlled variation of amplitude; (2) the amplitude distribution in the aperture is known and we wish to control the phase function to produce a shaped beam.

Case 1

The diffraction integral may be written

$$E(\theta)=f(\theta)\int_{-a/2}^{a/2}A(x)\exp[ikx(\cos\theta-\text{const})]dx. \tag{4}$$

Let $(ka/2)(\cos\theta-\text{const})$ be the angle variable u, and

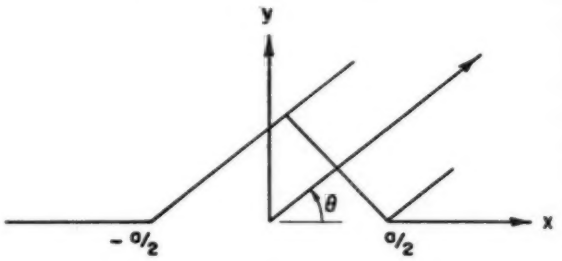


FIG. 3.

set $2x/a=\xi$, so that

$$E(\theta)=\tfrac{1}{2}af(\theta)\int_{-1}^{1}A(\xi)e^{iu\xi}d\xi. \tag{5}$$

Now the amplitude distribution $A(\xi)$ is expanded in a power series

$$A(\xi)=\sum c_n\xi^n \tag{6}$$

and Eq. (5) becomes

$$E(\theta)=f(\theta)\sum(-i)^n c_n[d^n g(u)/du^n], \tag{7}$$

since it may be observed that

$$d^n g(u)/du^n=\tfrac{1}{2}a\int_{-1}^{1}i^n\xi^n e^{iu\xi}d\xi,$$

in which we have set $g(u)=a\sin u/u$.

This result, Eq. (7), shows that for an even amplitude function in the aperture, the odd derivatives vanish, since the odd terms of the power series are zero, and the resulting diffraction pattern is a real function. Alternate even derivatives are of opposite sign, so that for a symmetrically tapered amplitude the summation

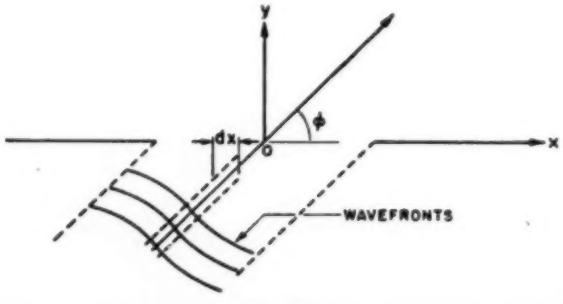


FIG. 2.

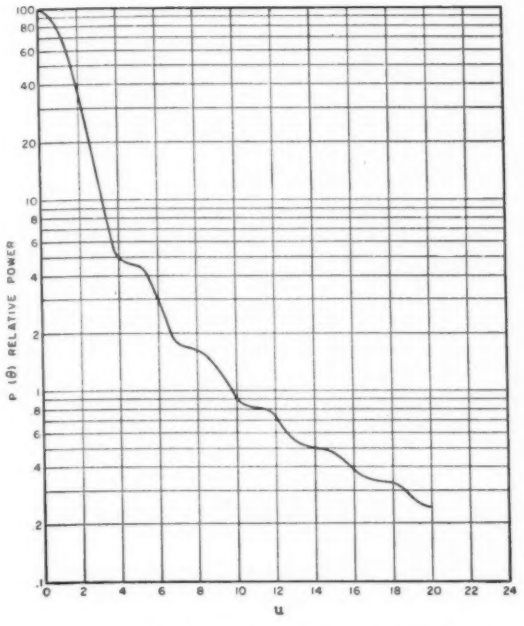


FIG. 4. Relative power pattern for triangular amplitude distribution.

of the terms of Eq. (7) leads to a diffraction pattern characterized by reduced secondary lobes.

If, however, the amplitude distribution is an asymmetric function, the resulting diffraction pattern is seen to be complex. Consequently, the pattern can be expressed by

$$E(\theta)=U_1(\theta)+iU_2(\theta), \tag{8}$$

where

$$U_1(\theta)=\sum(-1)^{l/2}c_l d^l g(u)/du^l; \qquad l=0, 2, 4, \cdots$$
$$U_2(\theta)=\sum(-1)^{\frac{1}{2}(m+1)}d^m g(u)/du^m; \quad m=1, 3, 5, \cdots.$$

The power pattern $P(\theta)$ is expressed by

$$P(\theta)=|E(\theta)|^2=U_1^2(\theta)+U_2^2(\theta). \tag{9}$$

It will be observed that the odd derivatives reach maxima at every point in which the even derivatives are zero; likewise, the even derivatives have maxima where the odd derivatives are zero. Therefore, the pattern $P(\theta)$ can have no zeros and will be characterized by a distribution of energy in a shaped beam.

As an illustration, consider the half-gable or triangular amplitude distribution $A(\xi)=1-\xi$. With the aid of Eq. (7) we write immediately

$$E(\theta)=f(\theta)[g(u)-idg(u)/du], \tag{10}$$

$$P(\theta)=f^2(\theta)\{g^2(u)+[dg(u)/du]^2\}. \tag{11}$$

This pattern is shown plotted as a function of the angle variable u in Fig. 4, for $f(\theta)=1$. Plotted in Fig. 5 are two other interesting amplitude distributions:[9]

$$\text{(I)} \quad A(\xi)=1-\sin(\tfrac{1}{2}\pi\xi),$$

$$P(\theta)=\frac{\sin^2 u}{u^2}+\frac{u^2\cos^2 u}{(\frac{1}{4}\pi^2-u^2)^2};$$

$$\text{(II)} \quad A(\xi)=\cos[\tfrac{1}{4}\pi(1+\xi)],$$

$$P(\theta)=\frac{\frac{1}{4}\pi^2+u^2-2\pi u\sin 2u}{(\frac{1}{4}\pi^2-4u^2)^2}.$$

The synthesis of a desired pattern by this method of controlled amplitude variation and uniform phase is limited by the practical consideration of obtaining the particular amplitude distribution. Some experimental results in support of the above theoretical development will be given in a later section.

Case 2

The amplitude distribution $A(x)$ is taken to be an even function, although there is no necessity for so choosing $A(x)$. The form of $A(x)$, however, is assumed known. Introducing the angle variable u as before, although in this case we have no constant phase term, we may write the diffraction integral

$$E(\theta)=\tfrac{1}{2}af(\theta)\int_{-1}^{1}A(\xi)\exp[iu\xi-i\psi(\xi)]d\xi. \tag{12}$$

Expansion of the exponential $e^{i\psi(\xi)}$ within the integral

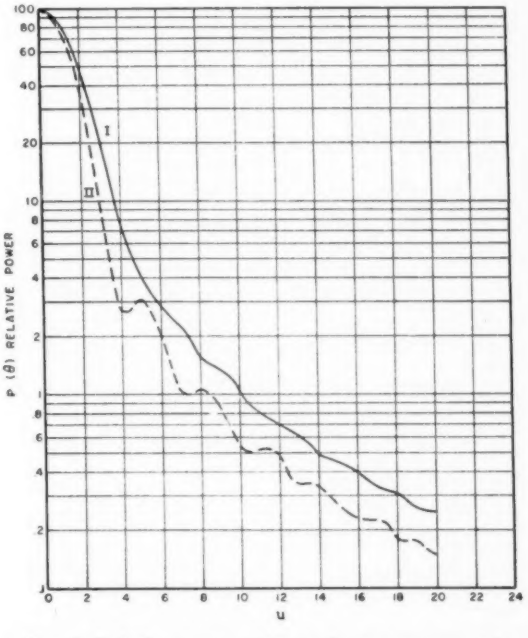


FIG. 5. Relative power patterns for two asymmetric amplitude distributions.

[9] A. S. Dunbar, Technical Report 13, Contract AF 19(122)78, Stanford Research Institute, Stanford, California (1950).

yields

$$E(\theta)=\tfrac{1}{2}af(\theta)\sum_{p=0}^{\infty}\int_{-1}^{1}A(\xi)e^{iu\xi}[(-i\psi)^p/p!]d\xi. \quad (13)$$

Because the first term of the expansion of the exponential is 1, we may write, as in Case 1,

$$E(\theta)=f(\theta)\sum(-i)^n c_n[d^n g(u)/du^n]$$

$$+\tfrac{1}{2}af(\theta)\sum_{p=1}^{\infty}\int_{-1}^{1}A(\xi)e^{iu\xi}[(-i\psi)^p/p!]d\xi. \quad (14)$$

The first term on the right of Eq. (14) is the expression for the diffraction pattern of an aperture with uniform phase. The second term is the modification of this pattern due to the variable phase. The resultant diffraction pattern is therefore the sum of two terms, either or both of which may be real or complex, depending upon both $A(\xi)$ and $\psi(\xi)$.

The synthesis of a shaped beam by the controlled variation of phase in an aperture having a known amplitude distribution may be accomplished by means of a method, stated by Chu,[4] based on the fundamental concept of the principle of the conservation of energy. In Chu's original method, which was applied to a cylindrical reflector and a line source, the shape of the reflector is calculated to produce the required shaped pattern from the given primary radiation pattern of the line source. This method has also been applied to the calculation of lenses and doubly curved reflectors to transform a given primary radiation pattern to the required shaped beam. By means of the principle of stationary phase this method may be stated in general form. Application will be made to the calculation and design of progressive-phase antennas.

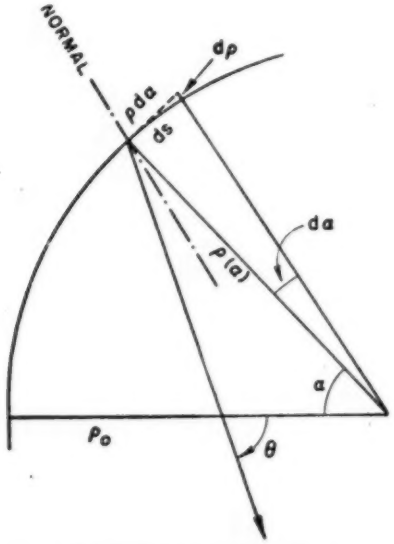


FIG. 6. Cylindrical reflector.

IV. THE GENERAL FORMULATION OF CHU'S METHOD

To express this method in general form, we postulate an amplitude distribution $A(s)$ on a cylindrical surface, a plane curve of which must be slowly varying and is restricted such that the curvature is always of the same sign. The far-field pattern is given by

$$E(\theta)=\int_{s_1}^{s_2}f(\theta,s)A(s)\exp[ik\varphi(s,\theta)]ds, \quad (15)$$

where $\varphi(s,\theta)$ is the phase function at the point P in the far field, and $f(\theta,s)$ is the form factor, which depends upon the polarization. Evaluation of Eq. (15) by the method of stationary phase yields the following results:

$$d\varphi(s,\theta)/ds=0 \quad (16)$$

and

$$E(\theta)=f(\theta,s_0)A(s_0)\exp[ik\varphi(s_0,\theta)]\times[-2\pi/\varphi''(s_0,\theta)]^{\frac{1}{2}}, \quad (17)$$

in which s_0 is the point at which the phase is stationary and φ'' denotes the second derivative with respect to s. Now Eq. (16) implies that s_0 is a function of θ; hence,

$$\varphi''(s_0,\theta)=(\partial^2\varphi/\partial s_0\partial\theta)d\theta/ds_0. \quad (18)$$

Therefore, Eq. (17) becomes

$$E(\theta)=f(\theta,s_0)A(s_0)\exp[ik\varphi(s_0,\theta)]\times[-2\pi/(\partial^2\varphi/\partial s_0\partial\theta)(d\theta/ds_0)]^{\frac{1}{2}}. \quad (19)$$

Consequently, since $P(\theta)=|E(\theta)|^2$, we may write

$$-(\partial^2\varphi/\partial s\partial\theta)P(\theta)d\theta=2\pi f^2(s,\theta)A^2(s)ds. \quad (20)$$

Equations (16) and (20) form the basis for the method of calculation to obtain an antenna that will produce a specified far-field pattern from a given amplitude distribution $A(s)$. Equation (20) relates a point on the antenna to a direction in the far field, and Eq. (16) provides a differential equation that must be solved for the required function $f(s)$ to produce the necessary phase function $\varphi(s,\theta)$.

In order to demonstrate this method, let us apply it to the cylindrical reflector, the geometry of which is shown in Fig. 6. It is required to calculate the curve $\rho(\alpha)$ which will transform the given primary radiation pattern $I(\alpha)$ to the specified pattern $P(\theta)$. We choose polarization perpendicular to the plane of the curve, so that the form factor becomes $\frac{1}{2}(\alpha-\theta)$. Let $\frac{1}{2}(\alpha-\theta)=\beta$. Then the phase factor is

$$\varphi(\alpha,\theta)=(2\pi/\lambda)[-\rho_0(1+\cos\theta)+2\rho\cos^2\beta],$$

and the diffraction equation, in which we have set $ds=\rho\sec\beta d\alpha$, becomes

$$E(\theta)=\int_{\alpha_1}^{\alpha_2}\cos\beta[I(\alpha)/\rho]^{\frac{1}{2}}\exp\{ik[-\rho_0(1+\cos\theta)+2\rho\cos^2\beta]\}\rho\sec\beta d\alpha$$

$$=\int_{\alpha_1}^{\alpha_2}[\rho I(\alpha)]^{\frac{1}{2}}\exp\{ik[-\rho_0(1+\cos\theta)+2\rho\cos^2\beta]\}d\alpha. \quad (21)$$

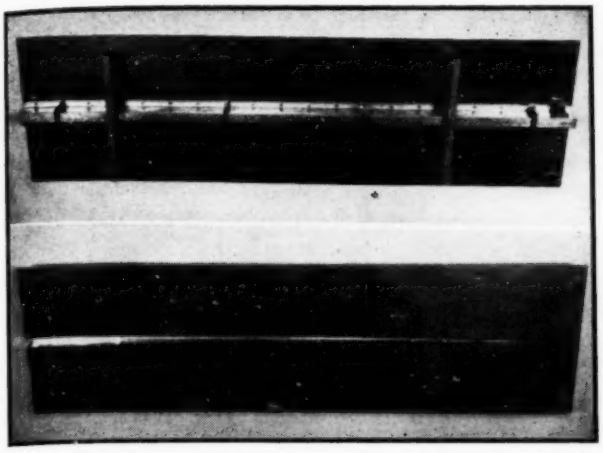


FIG. 7. Channel antenna for triangular amplitude distribution.

Performing the indicated operations on Eq. (21), we obtain the two equations

$$d\varphi(\alpha, \theta)/d\alpha = -2\rho \cos\beta \sin\beta + 2(d\rho/d\alpha)\cos^2\beta = 0 \quad (22)$$

and

$$P(\theta)d\theta = 2\pi I(\alpha)d\alpha. \quad (23)$$

Reduction of Eq. (22) yields $d\rho/\rho d\alpha = \tan\beta$, which is the differential equation to be solved for $\rho(\alpha)$. The solution is

$$\log(\rho/\rho_0) = \int_{\alpha_1}^{\alpha_2} \tan\beta d\alpha. \quad (24)$$

The foregoing illustrates that the general formulation applied to the cylindrical reflector problem yields the same results as obtained in the original formulation of the method by Chu. We now apply the method to a new problem, that of the progressive-phase antenna for beam shaping. The diffraction equation for the progressive-phase antenna is given by Eq. (3), i.e.,

$$E(\theta) = \int_{-a/2}^{a/2} f(\theta)A(x)\exp\{ik[x\cos\theta - \psi(x)]\}dx,$$

in which

$$\psi(x) = \int_{-a/2}^{x} \mu(x)dx; \quad \mu(x) = \lambda/\lambda_g(x) = c/v(x).$$

With the use of Eqs. (16) and (20), we obtain

$$d\psi(x)/dx = \cos\theta \quad (25)$$

and

$$\sin\theta P(\theta)d\theta = 2\pi f^2(\theta)A^2(x)dx. \quad (26)$$

The latter of these is the energy equation that relates a point on the progressive-phase antenna with a direction in the far field. It provides the functional relation

$$\theta = F(x).$$

Equation (25) is the differential equation to be solved for $\mu(x)$. The solution is simply

$$\mu(x) = \cos\theta. \quad (27)$$

Therefore, by adjusting the parameters of the antenna such that the phase velocity in the aperture satisfies Eq. (27) at each point in the aperture as required by Eq. (26), it is possible to design a progressive-phase antenna that will radiate a shaped beam. The practical application of this procedure of calculation to an experimental progressive-phase antenna will be described.

V. EXPERIMENTAL RESULTS

Beam shaping by controlled variation of the amplitude distribution, Case 1, has been demonstrated by means of long-slot or channel-guide antennas. This type of antenna has been described in the literature, and it has been demonstrated that the amplitude distribution may be shaped by tapering the slot of the channel-guide antenna.[10]

Figure 7 shows an experimental antenna consisting of a channel-guide in which the radiating slot is tapered in width for control of amplitude. The design was intended to produce a triangular or half-gable distribution. The measured amplitude distribution is plotted in Fig. 8.

A comparison of the measured and computed radiation patterns is shown in Fig. 9. The theoretical pattern is given by Eq. (11); the form factor $f(\theta) = \sin\theta$ for polarization in the plane of the long slot and perpendicular to its longitudinal axis. The angle variable $u = (ka/2)(\cos\theta - 0.9)$, since the ratio of the wave number of the wave in the slot to the free space wave number was determined experimentally to be 0.9. This pattern closely approximates the cosecant-squared curve.

Another long-slot antenna in which the slot width is constant was constructed and tested at the Ohio State University Research Foundation.[11] The amplitude distribution is plotted in Fig. 10. The resulting antenna

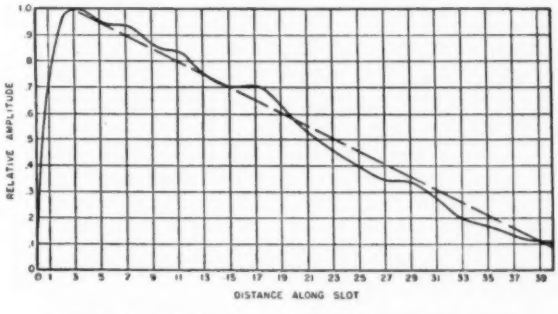


FIG. 8. Amplitude distribution of channel antenna.

[10] Ohio State University Research Foundation, Contract W33-038-ac-16520, e.g., Report 301-15, dated 1 May 1949; W. Rotman, Air Force Cambridge Research Center, Report E 5054, Jan. 1950.

[11] Published by permission of Ohio State University Research Foundation and the Air Material Command, Wright-Patterson Air Force Base, sponsor of Contract W33-038-ac-16520.

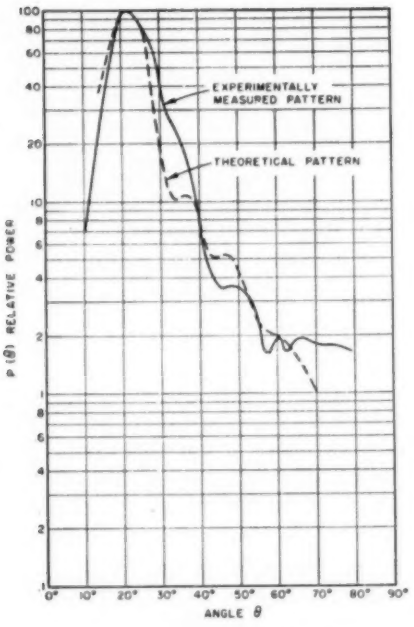


FIG. 9. Pattern of channel antenna.

pattern is shown in Fig. 11. This pattern approximates a cosecant-squared curve also.

The application of Case 2, the controlled variation of phase, to the design of a practical antenna has been demonstrated by means of a wave-guide slot array. The assumed amplitude distribution $A(x)=$constant; the specified far-field pattern $P(\theta)=\csc^2\theta\cos\theta$, $10°<\theta<50°$. Substitution of these into Eq. (26), noting that by choice of polarization, $F(\theta)=1$, yields, after integration

$$\int_{50°}^{\theta}\operatorname{ctn}\theta d\theta \Big/ \int_{50°}^{10°}\operatorname{ctn}\theta d\theta = x/l, \tag{28}$$

and the resulting functional relation between θ and x is

$$\ln\sin\theta=\ln\sin 50°+(x/l)[\ln(\sin 10°/\sin 50°)]. \tag{29}$$

Equation (27) becomes, therefore,

$$\mu(x)=\cos\{\sin^{-1}[\sin 50°(\sin 10°/\sin 50°)^{x/l}]\}. \tag{30}$$

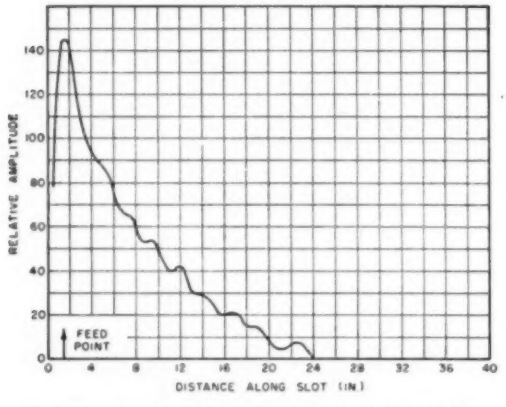


FIG. 10. Amplitude distribution of dielectric-filled channel antenna.

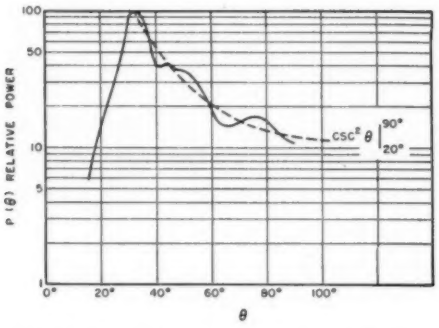


FIG. 11. Pattern of dielectric-filled channel antenna.

The controlled variation of $\mu(x)=\lambda/\lambda_g(x)$ in the wave guide was accomplished by the use of ridge wave guide, as shown in Fig. 12. The slots of the array are probe fed by magnetic probes, as illustrated in the figure. A photograph of the slot array is shown in Fig. 13. Inserts show details of the slots and the probes. The amplitude of the field on the array was adjusted by individual adjustment of the probes, so that the excitation of each slot is the same. Consequently, the probes are adjusted to couple successively greater and greater amounts of energy from the wave guide, as illustrated by the inserts of Fig. 13, in which the angle of the probe tip is an indication of the coupling. The experimentally measured and theoretically computed patterns of this antenna are shown in Fig. 14, with the $\csc^2\theta\cos\theta$ curve shown for comparison. The computed pattern represents a satisfactory fit to the specified curve. The experimental pattern, while not a good fit, is very encouraging and demonstrates the possibility of designing a shaped-beam antenna by means of controlled variation of phase. The discrepancies are attributed mostly to mechanical tolerances in the design of the ridge wave guide.

VI. ACKNOWLEDGMENTS

The author acknowledges the contributions and assistance of several individuals in the work described in this paper. Correspondence and discussions with J. E.

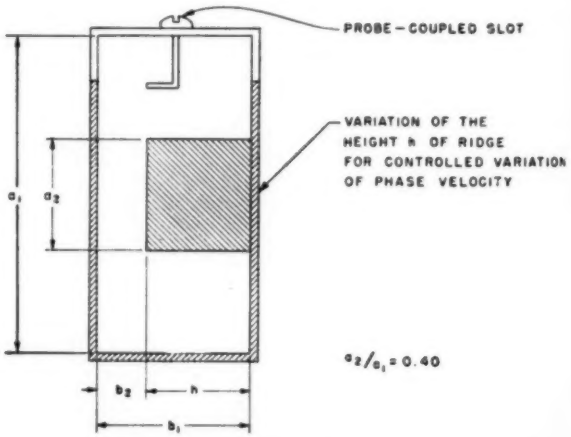


FIG. 12. Section of slot array.

Eaton, S. Silver, W. S. Lucke, and C. Flammer have been particularly helpful and stimulating. The data on the antenna credited to Ohio State University Research Foundation was supplied by C. H. Walter, whose cooperation is greatly appreciated. The computation of the theoretical pattern of the slot antenna was done by the Telecomputing Corporation of Burbank, California.

APPENDIX

The Curved Progressive-Phase Antenna

The conditions for the formulation of the calculation procedure for the curved progressive-phase antenna, in which the velocity $v(s)$ is constant, are contained in the general formulation. It is required merely to express the design equations in calculable form.

The phase function of a wave on a cylindrical surface, described in terms of the conventional rectangular

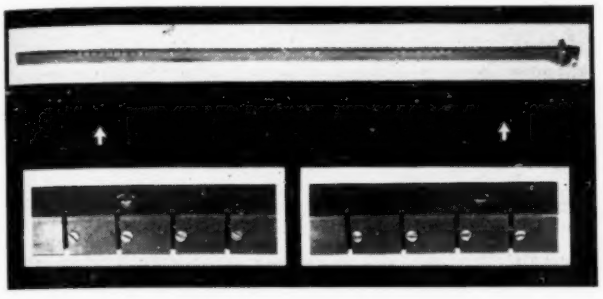


FIG. 13. Thirty-wavelength slot array.

coordinates and having no z dependence, is given by

$$\psi(s,\theta)=x\cos\theta+y\sin\theta-\int_{s_1}^{s_2}(c/v)ds. \qquad \text{(I-1)}$$

Now performing the operations required by the general formulation, we obtain first

$$d\psi(s,\theta)/ds=[\cos\theta+\sin\theta dy/dx-(c/v)ds/dx]dx/ds=0. \qquad \text{(I-2)}$$

Let $dy/dx=\tan\gamma$; then $ds/dx=[1+\tan^2\gamma]^{\frac{1}{2}}=\sec\gamma$; and

$$\cos\theta+\sin\theta dy/dx-(c/v)ds/dx=\cos\theta+\sin\theta\tan\gamma-(c/v)\sec\gamma=0, \qquad \text{(I-3)}$$

which yields

$$\cos(\theta-\gamma)=c/v,$$
$$\gamma=\theta-\cos^{-1}(c/v). \qquad \text{(I-4)}$$

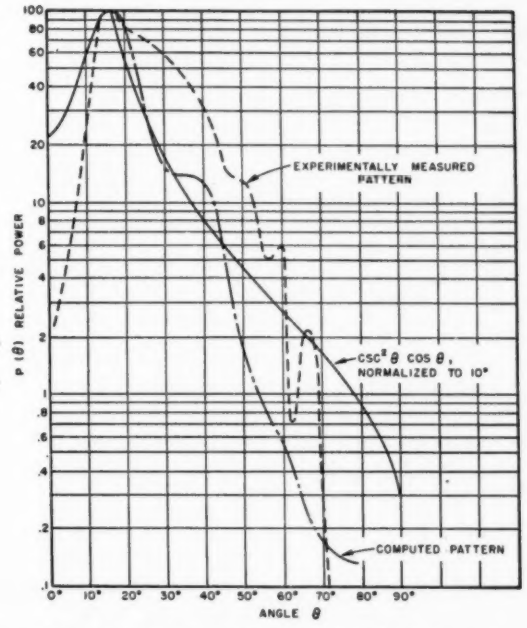


FIG. 14. Pattern of thirty-wavelength slot array.

Therefore,

$$y=\int_{x_1}^{x}\tan[\theta-\cos^{-1}(c/v)]dx. \qquad \text{(I-5)}$$

Secondly, we have

$$\frac{\partial^2\psi}{\partial s\partial\theta}=[-\sin\theta+\cos\theta\tan\gamma]\cos\gamma$$
$$=-\sin(\theta-\gamma)=-[1-(c/v)^2]^{\frac{1}{2}}. \qquad \text{(I-6)}$$

Therefore, the energy equation becomes

$$\cos[\theta-\cos^{-1}(c/v)][1-(c/v)^2]^{\frac{1}{2}}p(\theta)d\theta=f^2(x,\theta)A^2(x)dx. \qquad \text{(I-7)}$$

Note that for vertical polarization $f(x,\theta)\cong 1$ for $\theta>\gamma$, and for horizontal polarization $f(x,\theta)=\sin(\theta-\gamma)$.

It should be remarked also that (a) if $c/v=1$, Eq. (I-7) becomes indeterminate, and (b) if $(c/v)>1$, Eq. (I-7) becomes imaginary. This implies, of course, that the approximations of geometrical optics are not valid for a bound wave on the surface.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 8 AUGUST, 1952

A Point Focusing X-Ray Monochromator for the Study of Low Angle Diffraction*

LEON SHENFIL,† WARREN E. DANIELSON, AND JESSE W. M. DUMOND
California Institute of Technology, Pasadena, California

(Received February 18, 1952)

A point focusing x-ray monochromator was designed and constructed for low angle scattering studies. The anastigmatic point focus is achieved by means of two cylindrically bent quartz crystals whose focal circles are mutually perpendicular. The beam, emanating from the copper target of an x-ray tube, is reflected in succession, first from the crystal defining the horizontal focal circle, and second from the crystal defining the vertical focal circle following which it comes to a monochromatic point focus of wavelength 1.537A ($CuK\alpha_1$). The sample to be studied is placed between the second crystal and the point focus, and the scattered beam is detected by means of a photographic film placed at the point focus, at right angles to the undeviated beam, the latter being suppressed by means of an absorber, or allowed to pass through a hole in the film.

Mathematical analysis, in which a ray was traced through the two-crystal system, revealed correctly the shape and size of the point focus, and the possibility of reducing the latter in size by stopping down the beam emerging from the target.

INTRODUCTION

UP to the present, most low angle x-ray scattering work has been done with x-ray beams monochromatized, if at all, only by means of filters and collimated by means of pinhole or slit systems. An exception to this is the work of Guinier in France,[1] who with collaborators has used curved crystals to monochromatize and refocus an x-ray beam to a line focus. Guinier, following a suggestion by one of the present authors, has even tried two successive crystal reflections, with their focal circles in the same plane, in an effort to reduce the fogging due to incoherent crystal scattering. To our knowledge, however, one of us was the first to propose[2] a method of effecting two successive curved crystal reflections by means of elastically bent crystal laminas whose focal circles lie in mutually perpendicular planes, the geometry being such that the monochromatic radiation passed by the laminas converges to an anastigmatic point focus. The instrument here described was designed to exploit the possibilities of this proposal.

A list of references to recent work in the field of low angle x-ray diffraction is appended herewith.[3–6]

An endeavor has been made to design the present instrument so as to give high resolution in the low angle scattering range (down to 0.002 radians scattering angle) with a higher total intensity than has been achieved by the use of pinhole systems with comparable resolution. To accomplish this, use has been made of the focusing properties of two thin quartz (crystal) laminas, each of which is bent elastically to a cylindrical shape. These two laminas, which by two successive reflections focus the beam to a monochromatic point, constitute the essential feature of the instrument.

The advantages of a point-focusing monochromator are twofold:

(1) The focal spot plays a similar role in delineating the x-ray diffraction pattern to that of the cathode-ray spot in producing a picture on a television screen. Thus, when the focal spot is small in comparison to the interesting detail of the pattern, the interpretation of this pattern and the reduction and correction of the data are greatly facilitated.‡

(2) The incoherent radiation scattered by the first crystal (which is illuminated by the entire continuous x-ray spectrum) is much more effectively reduced by the second crystal than is the Bragg-reflected line intensity. Hence the focal plane is relatively freer from unwanted diffuse radiation than in single-crystal or slit systems.

The positions of the two laminas relative to the source of x-rays are shown in horizontal plan and vertical elevation in Fig. 1. (A perspective line drawing is also shown in Fig. 3 below.) In the plan view of Fig. 1 one can observe the x-rays emanating from the point source S towards the first crystal (crystal A), the position of the latter being such as to reflect the $K\alpha_1$ line of Cu. In the absence of the second crystal (crystal B) the beam comes to a vertical line focus FF. The three points S, A, and F_0 define the horizontal focal circle with center at C. (The points A and F_0 refer to the centers of crystal A and of the line focus FF, respectively.) It can be shown that the reflected beam appears to come from the virtual line source $F'F'$

* This work was performed and financed under the joint sponsorship of the ONR and the AEC by contract with the California Institute of Technology.

† Now with the Aerojet Engineering Corporation, Azusa, California.

[1] A. Guinier, Ann. phys. **12**, 161 (1939); J. chim. phys. **40**, 133 (1943); Bull. soc. chim. Belges **57**, 286 (1948); A. Guinier and G. Fournet, Compt. rend. **226**, 656 (1948).

[2] J. W. M. DuMond, Rev. Sci. Instr. **21**, 188 (1950).

[3] R. S. Bear and O. E. A. Bolduan, J. Appl. Phys. **22**, 191 (1951); O. E. A. Bolduan and R. S. Bear, J. Appl. Phys. **20**, 983 (1949).

[4] K. L. Yudowitch, J. Appl. Phys. **22**, 214 (1951); **20**, 1232 (1949); **20**, 174 (1949).

[5] C. G. Shull and R. C. Roess, J. Appl. Phys. **18**, 295 (1947); R. C. Roess and C. G. Shull, J. Appl. Phys. **18**, 308 (1947).

[6] L. H. Lund and G H. Vineyard, J. Appl. Phys. **20**, 593 (1949).

‡ The size of the point can be decreased at will by stopping down the beam emerging from the target. Without such stops and with our present x-ray tube (Machlett) the point appears slightly elongated with dimensions 0.196 mm by 1.40 mm. (Distance of scattering sample to focal plane, 66 cm.)

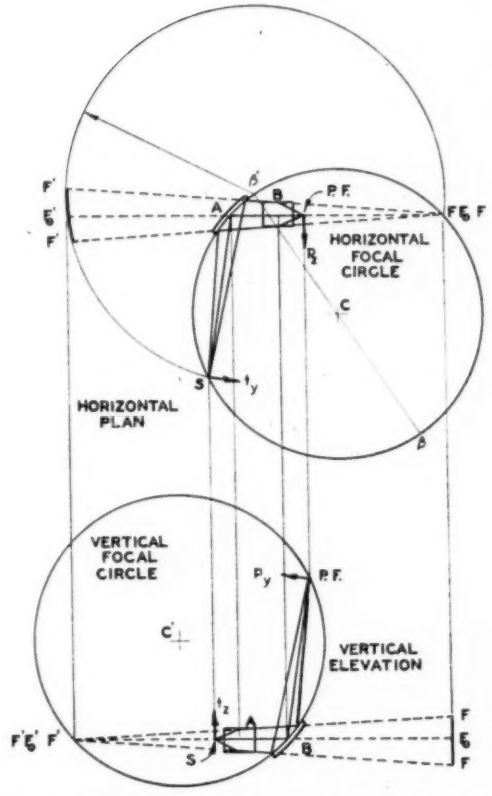


FIG. 1. Two bent crystals arranged to form a point focusing monochromator.

lying in the horizontal plane and having a radius of curvature $\beta'S$ as shown. (β is the common center of curvature of the bent atomic planes in crystal A, and β' is diametrically opposite β on the focal circle.) Thus, when the second crystal is in position, as is shown in the vertical elevation, the once reflected beam is reflected again to the point focus $P.F.$ The three points F_0', B, and $P.F.$ lie on a vertical focal circle with center at C' (where again the points F_0' and B refer to the centers of the virtual line source $F'F'$ and crystal B, respectively). If the latter were superimposed on the horizontal focal circle, then the above three points would coincide, respectively, with F_0, A, and S, the two focal circles being identical in size.

The sample is placed in the twice reflected beam near crystal B, and the photographic plate is placed at the point focus, at right angles to the central ray. The low angle diffraction pattern appears on the photographic film as a diffuse distribution of film grain in a small region about the direct beam, the latter being suppressed by means of an absorber, or allowed to pass through a hole in the film.

GEOMETRY OF A SINGLE BENT CRYSTAL

It is well known[7,8] that x-rays can be made to focus monochromatically to a line focus by a crystal lamina which has been hollowed out, cylindrically, to a given radius and then bent elastically to half that radius. Figure 2 shows a plan view of such a bent crystal with C the center of the focal circle, S a point source of x-rays, F the line focus, and β the vertical line (or generator) about which the crystal "planes" (cylinders after bending) are concentric. Lines $\overline{SP}$ and $\overline{PF}$ represent the incident and reflected rays (not necessarily in the plane of the focal circle). It is clear that if the arcs $S\beta$ and βF are equal, the angles of incidence and reflection $(\pi/2-\theta)$ are equal and constant for all positions of the point P, provided the vertical divergence angle (relative to the plane of the focal circle) is not excessive. The wavelength of the focused beam is given by the Bragg law $\lambda=2d\sin\theta$, where d is the grating spacing of the atomic reflecting planes. The position of point S, relative to the crystal, is such that the $K\alpha_1$ line of copper is reflected in the monochromator here described.

In Fig. 2, the point on the focal circle diametrically opposite β is denoted by β', and the center of the short piece of crystal lamina, ZZ', used in the present instrument is denoted by A. It should be clear from this figure that a much larger crystal lamina (such as that indicated, partly with lighter lines, between the points ZZ in the figure), with center at β' *could* be used to form the monochromatic line focus through F. The reason for using only a portion ZZ' entirely on one side of β' will be evident as soon as the two-crystal arrangement is discussed. The angle 2α is a measure of the angular displacement of point A from β'. α is also the angle between the atomic reflecting planes and the crystal face at A.

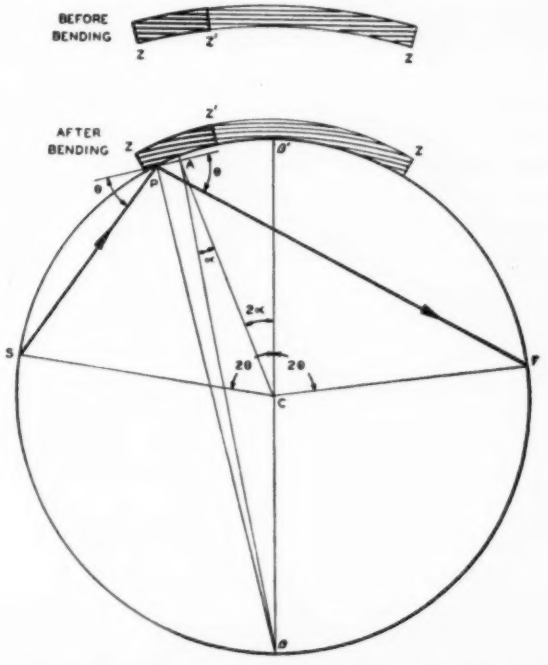


FIG. 2. Geometry of a single bent crystal, cylindrically ground before bending.

[7] J. W. M. DuMond and H. A. Kirkpatrick, Rev. Sci. Instr. **1**, 90 (1930).
[8] H. H. Johansson, Z. Physik **82**, 507 (1933).

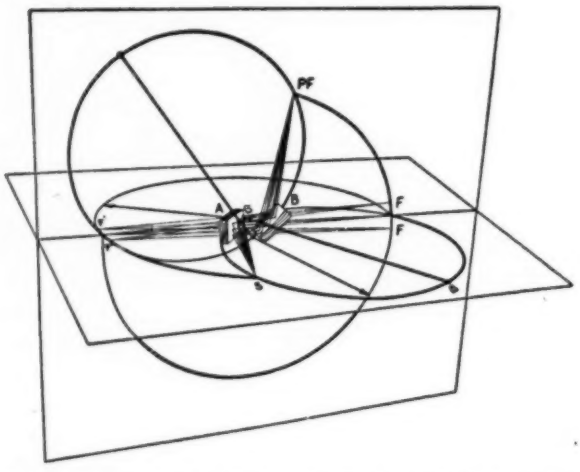


FIG. 3. Perspective line drawing to illustrate the 3-dimensional relationships involved in the geometry of the point focusing x-ray monochromator.

GEOMETRY OF THE POINT-FOCUSING MONOCHROMATOR

Now let us take a second identical crystal, as is shown in Fig. 1, with the plane of its focal circle perpendicular to the corresponding plane of the first crystal and place it in such a way that the line of intersection of the two planes is F_0F_0'. If we let $BF_0'=AF_0$ and if the point F_0' lies on the focal circle of the second crystal, then the rays striking this crystal will be reflected at the Bragg angle and will come to a point focus.

It should now be evident (1) why only a *part*, ZZ', of the total conceivable crystal ZZ shown in Fig. 2 is retained and (2) why the center A of the crystal ZZ' is displaced angularly by an amount 2α from the point β'. Both of these requirements originate from the fact that the first crystal must be kept entirely on one side of the point β' and the second crystal entirely on the other side so as to avoid mechanical interference between the two crystals or their crystal holders when the two focal circles are correctly mounted to avoid astigmatism. Clearly, the larger the angle α is made, the farther apart the two crystal clamping blocks may be placed. The distance between the centers A and B of the two crystals is given by $4r\sin\alpha\cos\theta$ where r is the radius of either focal circle.

Figure 3 is a perspective line drawing showing the two mutually perpendicular planes on which the two focal circles are described, together with the two curved crystals and the pencils of x-rays diverging from the source and reconverging to the anastigmatic point focus after the two successive reflections. Between the two crystals the beam is astigmatic, appearing to come from one line focus on the circular arc $F'F'$ and directed so that, were it not for the reflection at the second crystal, the beam would converge in the second line focus on the circular arc FF.

SIZE AND SHAPE OF THE POINT FOCUS

Since in practice one deals with a target of finite size and a spectral distribution ($K\alpha_1$ line) of finite width, it is reasonable to expect that the point focus will be of a finite size. To calculate this size, it is necessary to trace a general ray through the system. As a result of this analysis we shall find that only certain regions of the focal spot of the x-ray tube are so situated that the monochromator crystals can accept radiation from them within the limitations set by (1) the Bragg law at both crystals, and (2) the natural spectral breadth of the x-ray line. These regions, projected on a plane normal to the mid-ray of the beam from the source to the first crystal, we shall call "the usable projected target area." The size and shape of this area is imaged one-to-one in the focal spot formed after the two-crystal reflections.

In Fig. 4, S_0 represents the center of the target and S the point at which an arbitrary ray intersects the focal cylinder. Q is the point in which this arbitrary ray intersects the plane through S_0 normal to the mid-ray, $\overline{S_0A}$. The horizontal and vertical displacements of Q from S_0 are t_y and t_z, respectively.§ P is an arbitrary point on the crystal lamina, not necessarily in the plane of the focal circle. F_0 is the center of the line focus, and F is the point on the focal cylinder where the general ray SPF intersects it. $2t$ is the angle SCS_0 and -2η is the central angle ACP, both angles measured in the plane of the focal circle. (Since the focal circle is taken as having unit radius, $2t=SS_0$ so that t_y may also be thought of as the projection of the arc $2t$ on the t_yt_z plane.) The angle θ_0 satisfies the Bragg law $\lambda_0=2d\sin\theta_0$ where λ_0 is the center of the $CuK\alpha_1$ line. This angle is to be distinguished from θ_B, which is the Bragg angle for a wavelength λ falling at some arbitrary point within the natural breadth of the x-ray line whose peak or center is at λ_0. θ and θ' are the grazing angles of the general ray at the first and second crystals, respectively. Thus, if the (cylindrical) coordinates of the arbitrary point P are $1, 2\eta, z$, where z is the coordinate normal to the plane of the focal circle and the radius of the focal circle is reduced to unity for convenience, it is not hard to show that the relation between the grazing angle θ (angle between the general ray and crystal planes at P) and θ_B is given by

$$\theta-\theta_B=t-[z^2\tan\theta_0/8\sin^2(\theta_0-\alpha)]-[(\lambda-\lambda_0)/\lambda_0]\tan\theta_0, \quad (1)$$

where the quantities on the right are small compared to θ or θ_B. Furthermore, the grazing angle of the general ray at the second crystal is given by

$$\theta'-\theta_B=t'-[\eta^2/2]\tan\theta_0-[(\lambda-\lambda_0)/\lambda_0]\tan\theta_0, \quad (2)$$

§ The process of "reflection" in the first crystal keeps source and virtual image in the same horizontal plane, consequently the length of t_z is equal to that of its image. Hence, t_z also represents the vertical displacement of the image of Q from the mid-ray incident on the second crystal just as t_y is the horizontal displacement of Q from the mid-ray incident on the first crystal.

where $2t'=t_z/\sin(\theta_0+\alpha)$ is the angle in the vertical focal circle measured about C' as center from F_0' to the point of intersection of the general once-reflected ray with this circle (just as $2t$ in the horizontal focal circle is the angle about C from S_0 to S). The above two equations were derived for the case of cylindrically ground crystals prior to bending, henceforth to be called the exact focusing case. This is to be distinguished from the approximate or inexact focusing case in which one starts with flat crystals prior to bending.

In order that the general ray be reflected simultaneously from both crystals in succession, the following equation must be satisfied:

$$\theta-\theta_B=\theta'-\theta_B=0, \tag{3}$$

where the equality holds to within the diffraction pattern width ("dynamical" theory of Darwin or Ewald). This width in the case of CuKα_1 radiation and the (310) planes of quartz is 2 seconds of arc.

In Eq. (2) a number of approximations were made, namely the omission of terms of order higher than z^2, η^2, or ηz, where η and z, because of the limited size of the crystals, do not exceed 1/80 and 1/60 radians, respectively. The omitted terms are of the order of magnitude of the diffraction pattern width which is taken as the limit of negligibility.

Subtracting Eq. (2) from Eq. (1), we get the dependence of η and z on t and t' alone,

$$(\eta^2/2)\tan\theta_0-(z^2\tan\theta_0)/(8\sin^2(\theta_0-\alpha))=t'-t. \tag{4}$$

If the target of the x-ray tube is so small that t and t' can be neglected or if $t'-t=0$ in the above equation, the area of the crystal which is contributing to the point focus degenerates into two lines whose widths (3 mm) correspond to the diffraction pattern width and whose equations are

$$z=\pm 2\eta\sin(\theta_0-\alpha) \tag{5}$$

along which the wavelength decreases with increasing $|z|$, as is given by Eq. (1).

Since 2η and z can be considered as the two-dimensional coordinates of the point P relative to the center (A) of the first crystal, it is not hard to see that Eq. (4) represents two families of hyperbolas and Eqs. (5) represent their two asymptotes, with $t'-t$ playing the role of a parameter.

The extreme values of $t'-t$ occur when $\eta=\eta_{\max}=1/80$ rad, $z=0$ and when $z=z_{\max}=1/60$ rad, $\eta=0$. When substituted in Eq. (4) together with $\theta_0=40.7°$ and $\alpha=2.1°$, the range of $t'-t$ becomes

$$-0.04\leqslant t'-t\leqslant 0.04\text{ mm};$$

or, in terms of (t_y, t_z), coordinates of a point on the target projected onto a plane at right angles to the central incident ray (see Fig. 1)

$$-0.06\leqslant t_z-1.09t_y\leqslant 0.06\text{ mm}, \tag{6}$$

where use has been made of $t'=t_z/2\sin(\theta_0+\alpha)$ and $t=t_y/2\sin(\theta_0-\alpha)$.

The extreme values of t and t' depend on the fraction of the Kα_1 line which one wants to reflect. Thus, if the spectral width to be reflected is the full width at half-maximum for the Kα_1 line, which is 0.58 x.u. for Cu, then from Eqs. (1), (2), and (3),

$$\begin{aligned}&-0.10\leqslant t\leqslant 0.14\text{ mm}\quad\text{i.e.,}\quad -0.13\leqslant t_y\leqslant 0.17\text{ mm}\\&-0.10\leqslant t'\leqslant 0.15\text{ mm}\quad\text{i.e.,}\quad -0.14\leqslant t_z\leqslant 0.20\text{ mm}.\end{aligned} \tag{7}$$

From the inequalities (6) and (7), it is seen that the usable (projected) target area is a diagonal strip, approximately 0.09 mm wide and 0.45 mm long, inclined $47\frac{1}{2}°$ with the plane of the horizontal focal circle. For a 1-mm^2 target the usable strip is approximately 0.09 mm$\times$1.4 mm.

The general ray which emerges from point S with coordinates (t_y, t_z) will intersect the focal plane (plane through the ideal point focus and at right angles to the central ray) at a point with coordinates (p_y, p_z) (see Fig. 1), the coordinates of the ideal point being (0, 0). The relationships between the two sets of coordinates are

$$\begin{aligned}p_y&=t_z\sin(\theta_0-\alpha)/\sin(\theta_0+\alpha)\\p_z&=t_y\sin(\theta_0+\alpha)/\sin(\theta_0-\alpha),\end{aligned} \tag{8}$$

so that the inequalities to be imposed on p_y and p_z are almost the same as (6) and (7), namely,

$$\begin{aligned}&-0.06\leqslant p_z-1.09p_y\leqslant 0.06\text{ mm}\\&-0.13\leqslant p_y\leqslant 0.18\text{ mm}\\&-0.14\leqslant p_z\leqslant 0.18\text{ mm}.\end{aligned}$$

These inequalities indicate that the point focus is the image of the usable target area, namely, a strip 0.09 mm wide inclined at $47\frac{1}{2}°$ with the plane of the vertical focal circle. Furthermore, Eqs. (8) imply that it is possible to reduce the length of the point focus by stopping down the usable target area.

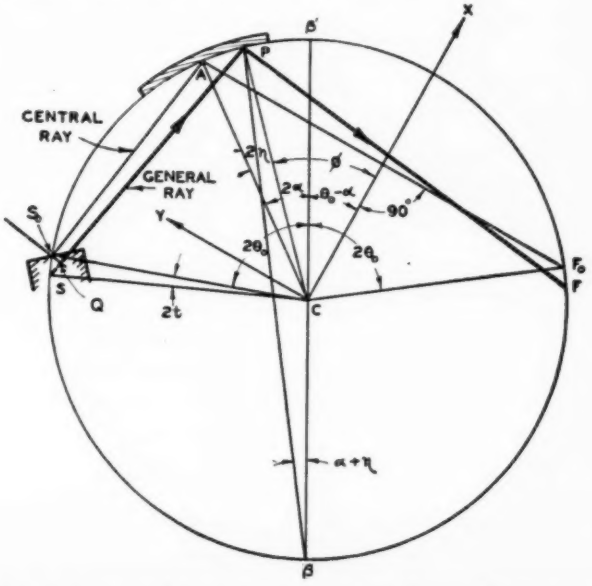


FIG. 4. Horizontal focal circle shown with its coordinate axes.

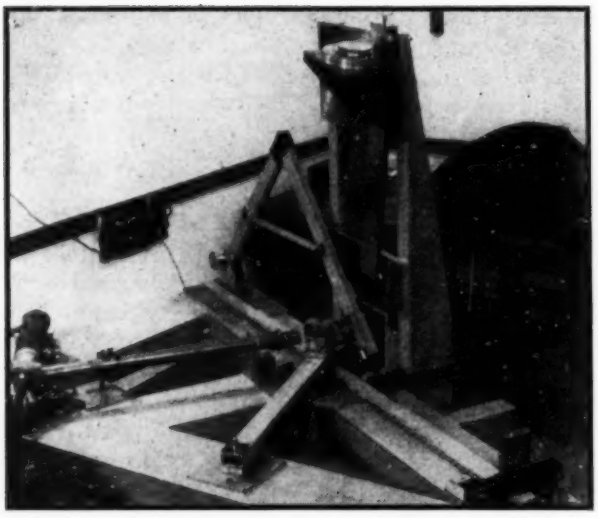


FIG. 5. The complete instrument with the exception of the sample holder and helium enclosure surrounding the two-crystal blocks.

An analysis, similar to the one above, carried out for the approximate focusing case, which is easier to realize in practice, shows that the point focus is again the image of the usable target area, namely, a diagonal strip, but in this case the dimensions of the strip are 0.19 mm×1.41 mm (for a 1-mm² target). Measurement of the exposed portion of a film placed in the focal plane of the monochromator for 2 minutes (the source of radiation being a Machlett tube operated at 35 pkv, 20 ma) shows it to be approximately a rectangular area 0.20×1.40 mm, a result which is consistent with the predicted value. This is at a distance of 75 cm from the second crystal or about 66 cm from the most distant position at which a scattering sample can be conveniently situated. Exposures of 20 or 30 hours, however, indicate that the size of the point focus is somewhat larger. This is probably due to the fact that the diffraction pattern of the quartz crystals extends over a wider angular range than was assumed in the above analysis.

DESCRIPTION OF THE POINT FOCUSING MONOCHROMATOR AND ITS OPERATION

The accompanying figure (Fig. 5) shows an over-all view of the instrument with the parts assembled for use, with the exception of the sample holder and the helium-filled housing surrounding the crystal clamping blocks and the x-ray paths. The x-ray tube is at the left and the two crystals may be seen, clamped between steel clamping blocks, at the center. A close-up view of the latter is shown in Fig. 6. The scattering sample is placed in the twice-reflected beam near the second crystal, and the film which records the diffraction pattern of the sample is held in place by a holder at the top of the vertical structure in such a way that the plane of the film is normal to the central ray and its center coincident with the point focus. With this arrangement, every point on a circle of specified radius (small in comparison to the sample-to-film distance) concentric with the central focal spot in the focal plane receives radiation which has been scattered under approximately the same angle. Thus aside from a negligible cosine correction arising from the fact that the rays in the direct beam are not parallel, the diffraction pattern represents truly the scattering property of the sample.

The cylindrical faces of the crystal clamping blocks (Fig. 6) are ground and lapped by a special method described by DuMond, Lind, and Cohen.[9] This shop method has the advantages that it automatically insures (1) that the generators of the precision cylindrical clamping surfaces shall be truly normal to plane ground reference faces on the top and bottom of the crystal clamping blocks, and (2) that the radii of curvature on blocks for crystals A and B shall be closely equal.

Since there are twelve degrees of freedom involved in lining up the two crystals relative to the x-ray tube, it is necessary to have a systematic procedure whereby each degree of freedom is taken care of independently. To accomplish this, use is made of the reference surfaces mentioned above as well as the reference surface provided on the triangular table. The line focusing property of a single bent crystal permits the alignment of each crystal (independently of the other) on its supporting arm. This is done (first with the "vertical" crystal assembly and then with the "horizontal" one) in the horizontal position, thus allowing the x-ray tube to remain stationary. After each crystal is placed correctly on its supporting arm, the crystals (with their arms) are then placed as shown in Fig. 5, and the

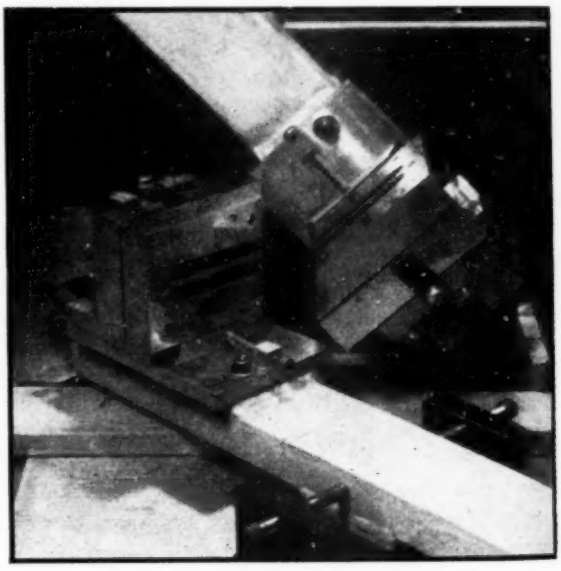


FIG. 6. Close-up view of the crystal blocks shown with the various adjusting screws.

[9] DuMond, Lind, and Cohen, Rev. Sci. Instr. **18**, 617 (1947).

instrument is corrected for astigmatism. This latter correction is made by pulling the second crystal toward away from the first along the line joining their centers.

The pertinent dimensions of the instrument are the following. The radius of curvature of the crystal laminas is approximately 120 cm; hence the radius of each focal circle is 60 cm. The maximum sample-to-film distance is about 66 cm, and the cross section of the beam upon emergence from the second crystal is 2 cm $\times$2 cm.

One of the present limitations of the instrument is the rather low intensity in the direct beam, which necessitates uncomfortably prolonged exposures for many scattering samples and becomes prohibitive in some cases. This difficulty arises from two effects. First, losses due to polarization occur because the two crystals behave in a manner similar to two Nicols which are crossed to almost complete extinction. (If the grazing angles θ and θ' were 45°, complete extinction would occur.)

Second, *not all* the monochromatic radiation reflected by the first crystal strikes the second crystal at exactly the Bragg angle. Hence, only a portion of this radiation is reflected by the second crystal. This loss is inherent in the instrument and cannot be corrected easily.

There are at least two ways of improving the intensity: (1) by replacing the present x-ray tube (a commercial model) with one whose specific intensity is 10 to 20 times higher; (2) by replacing the quartz crystals with topaz. The latter crystals have better reflecting properties than quartz, and, furthermore, because the grazing angle (for (006) or (303) planes and Cu$K\alpha_1$ radiation) is smaller than for quartz, the losses resulting from polarization are not as large.

A compensating advantage of the instrument is the very favorable ratio of intensity in the primary point-focused beam to the background intensity in the absence of a sample. This ratio is much better than in the case of a single-crystal monochromator. The diffuse background in a single-crystal monochromator comes chiefly from incoherent or diffuse scattering of the entire continuous spectrum incident on the crystal and also perhaps on parts of the crystal holder. This diffuse scattering is proportionately much more reduced by the second crystal reflection in the two-crystal point focusing monochromator than is the direct beam because the wavelength and direction of the rays are appropriate for the second Bragg reflection to a much greater extent in the once reflected beam than in the diffusely scattered beam.

Figure 7 is a photograph of the entire instrument with the working parts, including the entire x-ray path from tube to diffraction camera, enclosed in a gas-tight helium-filled housing, Since the x-ray path is 160 cm long a considerable loss of x-ray intensity through air absorption is thus eliminated. The vertical part of this housing (slightly inclined toward the observer) which encloses that part of the beam which has undergone the second crystal reflection can be clearly seen. The scattering sample, in the form of a thin layer between Nylon films, is secured to a brass plate and can be introduced into the beam at any one of four clearly visible points, one of which, near the bottom, is quite close to the second crystal. The distances from sample to photographic film which these four locations permit are 66, 36, 16, 6 cm. These distances have been very carefully laid out in machining the slots for introduction of the brass plates (like drawers) into the beam. The selection of the best sample-to-film distance is made as a compromise between angular resolving power and exposure time required. Large diffraction patterns at fairly wide angles which will, in consequence, not have very great surface intensity because the scattered energy is distributed over a wide area are to be placed at the closer sample-to-film distances.

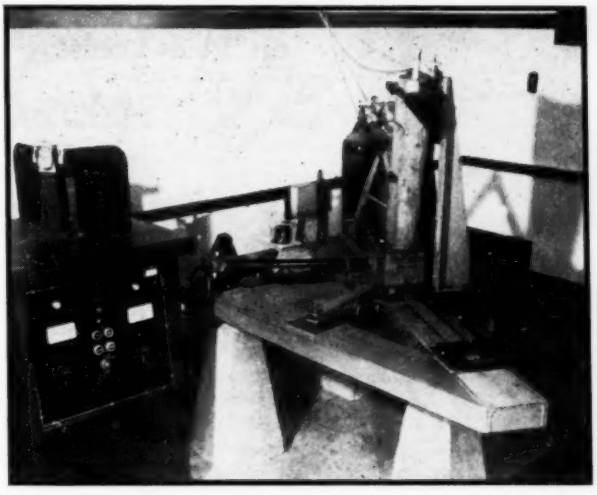


FIG. 7. The complete point focusing monochromator as used in low angle diffraction experiments.

In conclusion we wish to express our gratitude to the ONR for their generous support and sponsorship of the development of this instrument and to Professor Linus Pauling of this Institute for his faith in and support of the project in recommending it to its present sponsors when its success was still problematical. We are also most grateful for the friendly advice and cooperation of Professor R. B. Corey of the C.I.T. Chemistry Department and to Professors J. Weigle and G. W. Beadle of the C.I.T. Biology Department for most helpful advice and discussions regarding possible problems suitable for the new instrument.

Latex Particle Size Determination Using Diffraction Peaks Obtained with the Point Focusing X-Ray Monochromator*

WARREN E. DANIELSON, LEON SHENFIL,† AND JESSE W. M. DUMOND
California Institute of Technology, Pasadena, California

(Received February 18, 1952)

We describe the results of experiments we have made, using the point focusing monochromator as the primary tool, to determine the particle size of latex spheres. The suitability of the instrument for this particular study is described and the experimental data obtained are tabulated, these data coming from our experiments with three physically distinct samples of Dow latex. We attempt to make a critical interpretation of the data by considering separately several possible space arrangements which the latex spheres might assume when the water, which is initially the suspending fluid, is evaporated. Corrections for the finite size and shape of the "point" focus are described. The absence of a significant difference in the mean particle sizes of the three samples considered is established.

The data from all three samples is combined to yield a mean particle diameter under an external pressure of one atmosphere of 2687.5A with a statistical standard deviation of 1.2A and a fixed (systematic) error estimated to be not more than ±7A.

SUITABILITY OF THE POINT FOCUSING MONOCHROMATOR FOR THIS EXPERIMENT

A DETERMINATION of the mean particle size of latex spheres by observing the x-ray diffraction pattern obtained with these objects as scatterer was first carried out by Yudowitch[1] using slit-system collimation and approximate monochromatization by filtration. The widespread interest[2] in the particle size of the now famous Dow latex, batch 580-G, lot 3584 (extensively used as a valuable comparison standard of size for electron microscopes which have revealed that they consist of extremely uniform spherical particles about 2600A units in diameter) and the belief that we could increase the precision with which this size can be obtained from x-ray diffraction data have prompted us to make similar latex diffraction studies, but with the more strictly monochromatic sharply converging beam afforded by our newly developed point focusing x-ray monochromator.[3]

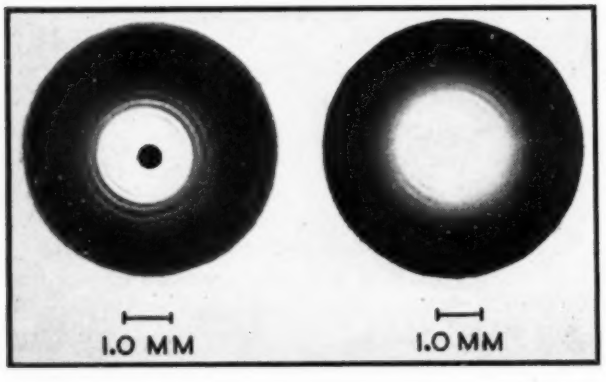


FIG. 1. The diffraction pattern reproduced on the left was obtained in 129.6 hours using sample I-*b* with the "fine" focus arrangement, while the one on the right was obtained in 91.1 hours using sample *II* with full intensity.

The monochromatization is such that only the $CuK\alpha_1$ line contributes radiation to the converging beam in which the sample being studied is placed. Furthermore, the size of the focus obtained can be conveniently varied from a nearly circular spot about 0.2 mm in diameter (at as much as 66 cm from the diffraction sample) to an elongated one about 0.2 by 1.2 mm, as dictated by the resolution desired, merely by introducing a slit in front of the x-ray tube. A third important characteristic of the beam produced by the point focusing monochromator is its virtual freedom from background, this desirable feature being attributable to the two successive Bragg reflections which the beam undergoes. The incoherent scattering of the continuous spectrum at the first crystal is much more effectively suppressed after the second reflection than the principal monochromatic beam.

EXPERIMENTAL RESULTS

Our sample of latex was taken from a 5-cc vial containing 40 percent water suspension of latex‡ (this was the standard container and contents sent to electron microscopists by the Dow Chemical Company). The water was removed by evaporation and about 0.4 cc of the latex powder was placed between 1-mil Nylon sheets in a cavity of dimensions 2×2×0.1 cm. This sample was then placed in the converging beam slightly above the second crystal of the point focusing monochromator with the 2×2-cm square faces normal to the central ray which passes through the centers of the squares. The sample-to-film distance is given below in conjunction with the other quantitative experimental data (the film is, of course, normal to the central ray

* This work was performed and financed under the joint sponsorship of the ONR and AEC by contract with the California Institute of Technology.

† Now with the Aerojet Engineering Corporation, Azusa, California.

[1] K. L. Yudowitch, J. Appl. Phys. **22**, 214 (1951).

[2] C. H. Gerould, J. Appl. Phys. **21**, 183 (1950).

[3] For description, see companion paper to this one, Shenfil, Danielson, and DuMond, J. Appl. Phys. **23**, 854 (1952).

‡ This sample was kindly loaned to us by Dr. R. F. Baker of the University of Southern California.

at the point focus with its center coinciding with the point focus). The observed diffraction pattern consists of a series of well-defined rings corresponding to intensity maxima for various (small) angles of deviation in the x-ray beam.

Figure 1 shows typical diffraction patterns obtained with "coarse" and "fine" adjustment of the focus, respectively. In the "coarse" adjustment the length of the slightly elongated focus is sufficient to blur out the successive diffraction rings in one azimuth as can be clearly seen, while in the "fine" adjustment the "point" focus is sufficiently short along its greatest dimension to make the rings distinguishable in all azimuths. The dark portion in the center of the left-hand picture is a hole in the film to permit the direct beam to pass freely through the film without undue fogging. Successive rings, which are seen to be clearly resolved, are separated by a difference of scattering angle of about 0.002 radian. Figure 2 shows a microphotometer trace of the 129.4-hour exposure made using high resolution.

In compiling our data we have used diffraction photographs taken with three physically distinct samples of the Dow latex particles. Two of these, which we shall call samples I-a and I-b, respectively, were from physically distinct portions of the latex loaned to us by Dr. Baker of U.S.C., and the third sample, which we shall call sample II, was latex from the same 5-cc vial as that used by K. L. Yudowitch[1] in his work on latex. This last sample was kindly loaned to us by him so that we could investigate the possibility of variation in the average particle size from distinct apportionments of the same Dow Chemical batch. Such a possibility suggested itself when we found a significant disagreement between the results of our measurements and those of Dr. Yudowitch.

Before discussing the conclusions we draw from our measurements, we present Table I containing the pertinent experimental data together with the inferred particle diameter obtained from calculations to be described below. The following symbols and numerical values are used:

λ = wavelength of $CuK\alpha_1$ line = 1.5374A.
D = particle diameter of latex spheres.
d = sample-to-film distance (varied from 64 to 66 cm).
ϵ = angle through which x-ray beam is deviated by scatterer.
$u = 2\pi D\epsilon/\lambda$.

Values for successive diffraction ring radii were measured from microphotometer traces such as the one shown in Fig. 2; the microphotometer "window" size for most of the traces corresponds to a film area of 0.06×0.13 mm. For the weaker, larger diameter, diffraction rings (rings beyond No. 11), it was found that the effect of film grain could be lessened, and consequently more reliable results obtained, if the "window" were lengthened and an experimental correction made for the systematic error so introduced.

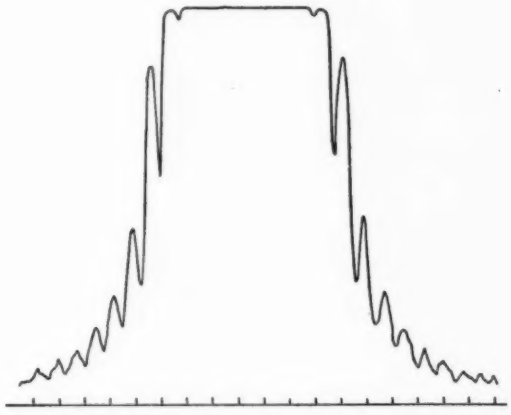


FIG. 2. Microphotometer curve diffraction pattern of latex particles (sample *I-b*, fine focus, 119.4-hr exposure). Intervals at the bottom represent 0.50 mm on the diffraction pattern or a scattering angle of about 3 minutes of arc.

INTERPRETATION OF DATA

In order to be able to infer particle sizes for the latex spheres from the entire series of rings, we have considered at least three possible space arrangements of the particles relative to each other in each of which the particle diameter might be expected to influence the diffraction pattern differently.

(1a) The spheres may tend to clump, in the process of drying, in such a way that the interior of each clump is made up of a close-packed hexagonal or cubic array of particles like a crystallite in a polycrystalline solid, the orientation of the clumps being random.

(1b) It is also conceivable that the spheres may tend to form a close-packed array that is a hybrid of the hexagonal close packing (layer scheme *A*, *B*, *A*, *B*, *A*, *B*, etc., see Fig. 3) and the cubic close packing (layer scheme *A*, *B*, *C*, *A*, *B*, *C*, etc.) such that the layers in which each sphere makes contact with six others are placed unsystematically upon one another instead of having a definite relation to the lower layers.

(2) The particles surrounding any arbitrarily chosen particle may fall into a spherically symmetric arrangement similar to that commonly assumed as representative of the disposition of atoms in a liquid, but with no other more far-reaching type of regularity markedly present.

(3) The particles may be distributed with sufficient lack of regularity that the distribution is essentially random.

One might suppose that electron microscope pictures of the latex such as those given by Gerould[4] offer strong support to the first possibility (1a). However, if crystal-type packing were the major cause for the observed diffraction rings (as in a Debye-Scherrer powder pattern), there would be a series of rings

[4] C. H. Gerould, J. Appl. Phys. **21**, 185 (1950).

TABLE I.

Peak number	ϵ_{hkl} (milli-radians)			D_{hkl} (A) ($P=0$, not corrected)		
	I-a	I-b	II	I-a	I-b	II
5	2.8018	2.8218	2.8312	2709.9	2690.7	2681.8
6	3.3825	3.3780	3.3855	2703.9	2707.5	2701.5
7	3.9520	3.9829	3.9792	2706.2	2685.2	2687.7
8	4.4987	4.5441	4.5036	2721.0	2693.8	2718.0
9	5.1286	5.1376	5.0912	2687.8	2683.1	2707.6
10	5.6975	5.6927	5.7107	2690.1	2692.3	2683.9
11	6.2700	6.2490	6.2817	2690.4	2699.4	2685.4
12	6.8341	6.8432	6.8081	2693.8	2690.2	2704.1
13	7.4111	7.4097	7.3994	2691.9	2692.4	2696.1
14	7.9961	8.0064		2687.5	2684.1	
15	8.5714			2683.7		
16	9.1529			2684.2		
17	9.7075			2689.4		

corresponding to each fundamental spacing of "crystallite" planes containing a high surface density of particles. Hence the fact that we observe only a single distinct series of rings seems to be proof that such scattering is at most a minor contributor to the diffraction pattern. The other possibility (1b) of a hybrid "crystal" built up as outlined above would indeed lead to clumps with only one fundamental set of interplanar spacings. However, we find that such an interpretation of the results yields a value for the particle size that is wholly incompatible with electron microscope values, and furthermore, the relative intensities between successive rings as observed do not correspond to those which would be expected in such a case.

The development of a rather idealized theory to describe the space arrangement of the particles as conceived under (2) above was given by Gingrich and Warren[5] in 1934 and has been applied to the present problem by K. L. Yudowitch.[1] The main simplifying assumption underlying the theory is that the number of spheres per unit volume, as a function of the radial distance (r) from an arbitrarily chosen sphere, is essentially constant except for a peak at $r=D$ and a void for $r<D$. To describe the size of the peak at $r=D$, a "packing" parameter (P) is introduced. It is so defined that $P=0$ corresponds to a random distribution of spheres; hence case (3) considered above is included in this treatment. Denoting the x-ray intensity scattered at an angle ϵ by I and setting $M=$number of spheres in the sample, $N=$number of electrons per sphere, and $\Phi(u)=\frac{3}{u^3}(\sin u-u\cos u)$, the simplification mentioned above leads to the following formula:

$$I=MN^2\Phi^2(u)\left[1+P\left\{5\frac{\sin 2u}{2u}-6\Phi(2u)\right\}\right]. \quad (1)$$

$\Phi(u)$ is zero when $\tan u=u$ (except at $u=0$); hence I has an oscillatory character that would account for a series of intensity maxima. In so far as a formula of this type accurately describes the intensity pattern, measurements of the larger rings obtained yield a particle diameter that is more reliable than one made using smaller rings because they are less sensitive to changes in P.

A formula quite similar to Eq. (1) above, but based on Rodriquez' work[6] on the kinetic theory of fluids, has been given by Fournet,[7]

$$I=MN^2\Phi^2(u)[1+(8v_0\nu/v_1)\Phi(2u)]^{-1}, \quad (2)$$

wherein v_0 and v_1 are the true and mean particle volumes, respectively, and ν is nearly constant. The theoretical calculation of ν requires a knowledge of the interaction potential between the particles so that except for relatively few cases,[6] it must be determined experimentally. One of the main assumptions underlying (2) is the supposition that the probability distribution function describing the particle positions does not differ markedly from $e^{-\varphi(r)/kT}$, where $\varphi(r)$ is the interaction potential between spheres. For particles as large as the latex spheres under consideration, the validity of this assumption may be questioned. Nevertheless, for large values of u (or small values of P and $\nu v_0/v_1$), Eqs. (1) and (2) become identical.

TABLE II.

Peak number	u_{max} (radians) for $P=0$	u_{max} (radians) for $P=\frac{1}{2}$
5	15.515	15.602
6	18.689	18.761
7	21.854	21.915
8	25.013	25.066
9	28.168	28.215
10	31.320	31.362
11	34.471	34.509
12	37.619	37.654
13	40.767	40.799
14	43.914	43.944
15	47.060	47.088
16	50.206	50.235
17	53.351	53.375

We have based our size determinations upon Eq. (1). Table II gives the values of u for which the intensity has a relative maximum for the cases $P=0$ and $P=\frac{1}{2}$.

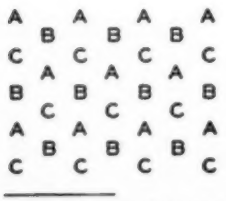


FIG. 3. To illustrate the close-packed arrays considered in the text. Any set of like letters is to be thought of as representing the centers of spheres in a given horizontal plane, these sets are then to be thought of as lying in different horizontal planes.

EFFECT OF FINITE SIZE OF POINT FOCUS AND MICROPHOTOMETER WINDOW

Probably the most important systematic error introduced into the above calculations, except possibly

[5] N. S. Gingrich and B. E. Warren, Phys. Rev. **46**, 248 (1934).

[6] A. Rodriquez, Proc. Roy. Soc. (London) **A196**, 73 (1949).

[7] G. Fournet, Acta Cryst. **4**, 293 (1951).

that due to inadequacies in the theory which yields Eq. (1), is the error which results from interpreting the data as though the primary beam converged to a mathematical point instead of the finite elongated spot observed experimentally. Before describing the correction for this error, it is advantageous to consider the qualitative features of the primary beam. Figure 4 shows two reproductions of exposures made, in the focal plane, with the primary x-ray beam. The 30-minute exposure shows the elongated nature of the point focus very clearly, while the longer exposure (31 hours) shows the magnitude and distribution of those much weaker parts of the primary radiation which are not focused in this elongated spot. All of the major features of the 31-hour exposure are readily explained: The streamer which makes an angle of about 45 degrees with the major axis of the focal spot is due to CuKα_1 radiation that has been scattered (coherently but diffusely, i.e., not at the Bragg angle) by the first crystal in such a fortunate direction that it is subsequently focused (in a line) by the second crystal. The small spot which lies along the major axis of the focal spot about 1½ mm from its center represents a "point focus" for the CuKα_2 line, the accompanying (very faint) streamer being due to CuKα_2 radiation scattered by the first crystal and focused by the second. It is important to note that this non-Bragg reflected radiation has been completely eliminated from the focal plane except in two well-defined regions where it can be readily distinguished from radiation scattered by the sample being studied.

The finite dimensions of the focal spot introduce an important systematic error in the diffraction ring diameters to a different degree for different azimuths of the pattern and different rings. The azimuth of best resolution normal to the long axis of the focal spot was the one invariably used for measuring ring diameters. Figure 5 shows a map of the distribution of x-ray intensity over the focal spot as it was used for most of the latex work. The spot was divided into six annular sections or zones in such a way that the arcs defining these sections are concentric with the point of maximum intensity for a specified diffraction ring taken on the azimuth of greatest resolution, and the total direct beam intensity in each zone is assumed to be concentrated at the mid-point of the mean arc in that zone. Hence the effect of the true beam has been approximated by six ideal beams coming to a small array of true point foci along the azimuth of best resolution, an array which would yield essentially the same intensity distribution in the neighborhood of a given ring as that actually obtained. Finally, the radial position of the maximum expected for scattering from these six ideal beams is compared with the position of the maximum expected for a single central ideal beam. The corrections to the particle size obtained by this procedure are shown in Fig. 6.

The size of the microphotometer window was such

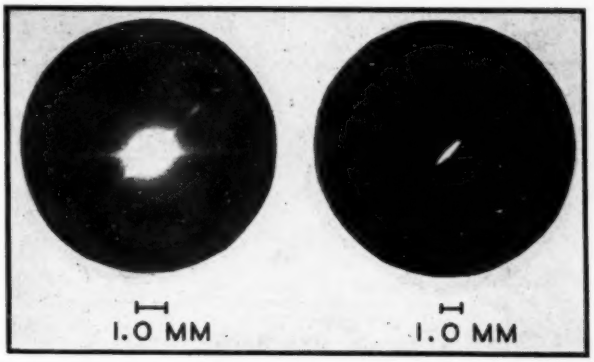


FIG. 4. On the left is shown a reproduction of a 31-hour exposure to the main beam at the "focal point," the major features of which are discussed in the text. The appearance of a film placed at the focal point and exposed for only ½ hour is shown at the right.

that no appreciable error is introduced in assuming it to be a point scanner. The corrections just obtained for the finite size of the focus are not included in Tables I and III, but are included in the final results, Table IV.

COMPARISON OF SAMPLES

The weighted average values of D together with the statistical standard deviations σ_i, for the case in which all measured rings for a given sample are taken as statistically independent are given in Table III.

We have included the results of separate calculations

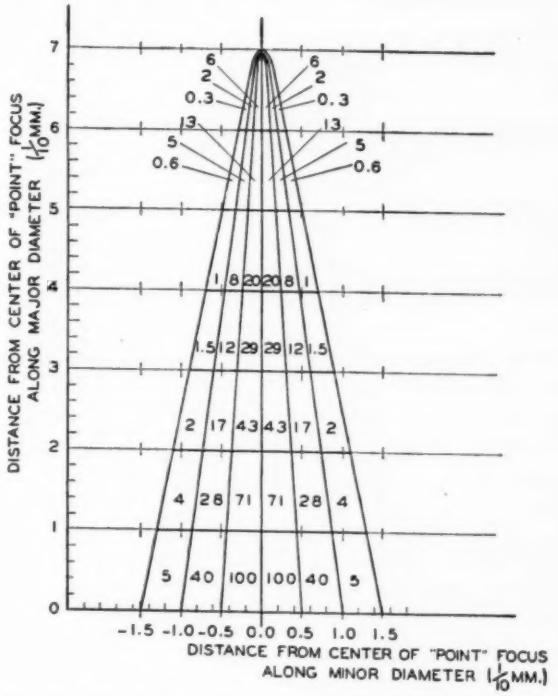

DISTANCE FROM CENTER OF "POINT" FOCUS ALONG MAJOR DIAMETER (1/10 MM.)
DISTANCE FROM CENTER OF "POINT" FOCUS ALONG MINOR DIAMETER (1/10 MM.)
-1.5 -1.0 -0.5 0.0 0.5 1.0 1.5
0 1 2 3 4 5 6 7
6 2 0.3 13 5 0.6
1 8 20 20 8 1
1.5 12 29 29 12 1.5
2 17 43 43 17 2
4 28 71 71 28 4
5 40 100 100 40 5

FIG. 5. This diagram shows the shape of half of the primary "point" focus, the other half being essentially the reflection of the part shown in the horizontal axis. The number in each subregion indicates the mean relative intensity there, as determined by microphotometer traces of a series of exposures made with the direct beam.

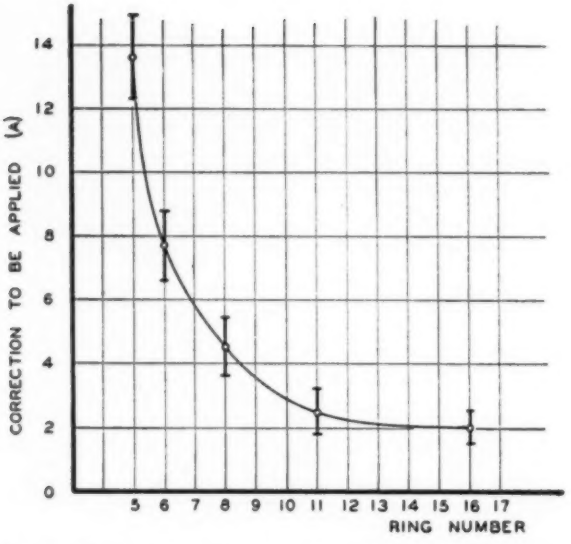


FIG. 6. Showing the correction to be applied to the diameter obtained from measurements of a given ring because of the finite size of the "point" focus.

for $P=0$ and $P=\frac{1}{2}$ because the standard deviations given *do not reflect the effect of systematic errors* such as those introduced by errors in the assumed theory. Other systematic errors such as those due to film shrinkage, sample-to-film distance, and microphotometer distance calibration are believed to be less than $2\frac{1}{2}$A. Although systematic deviations of relative ring diameters from their theoretically predicted values still remain in the results at the stage of Table III, the error introduced is essentially the same in the case of each sample measured, so that we can conclude at this point that *no significant difference in the mean particle size of the three samples used is indicated.*

Having established this result, we combine the data by ring numbers, considering the different measurements of a given ring (on all the different exposures in which it can be measured) *as statistically independent.* Table IV (using $P=0$) presents the results of this treatment. Here we attach an internal as well as an external precision index to the value of the particle diameter obtained from all measurements of each ring. This internal index for a given ring is proportional to the reciprocal of the square root of the total (summed) weight of the individual measurements of this ring; the external index is the statistical standard deviation from their mean of the measurements involved. In combining the results by rings to give a final weighted mean diameter, we have taken W, the weight, proportional to $[2/(\sigma_{int}^2+\sigma_{ext}^2)]^{\frac{1}{2}}$. Also, we have omitted data from peak number 8. This has been done as a result of definite evidence that this ring is *distorted due to radiation that is not scattered by the latex.* This distortion is, in fact, due to the CuKα_2 streamer referred to above in reference to Fig. 4. Figure 7 shows a plot of inferred particle diameter *versus* the number of the intensity maximum whose position was used for the calculation, both for $P=0$ and for $P=\frac{1}{2}$. In the case of $P=0$, there appears to be little, if any, suggestion of a decreasing diameter with higher order maxima; however, there does seem to be such a trend if calculations are based on a value of $P=\frac{1}{2}$.

We take the result based upon $P=0$, rather than $P=\frac{1}{2}$, as the more reliable because of the better external consistency obtained and estimate that, in view of the difference of 5.5A between the value based upon $P=0$ and that based upon $P=\frac{1}{2}$, the systematic error due to inadequacy of the theory is not likely to be more than about 7A.

USE OF RELATIVE INTENSITIES AT DIFFRACTION PEAKS

In all of the calculations made so far we have been concerned only with the positions of the intensity maxima. Since the pictures obtained also yield (1) the positions of the minima and (2) the relative intensity at various points in the pattern, we should consider how such data can be used to tell us more about the particle size. If the point focus were many times smaller than the distance between successive rings and a suitably small microphotometer opening were used,

TABLE III.

Sample	Diameter in A (not corrected) $P=0$	Diameter in A (not corrected) $P=\frac{1}{2}$	External precision index, σ_i (by individual samples)
I-a	2693.7	2697.6	2.0
I-b	2690.7	2695.9	2.4
II	2696.7	2701.7	2.7

the intensity should drop to very nearly zero when $\Phi(u)=0$ *independent of the packing.* However, we have been as yet unable to realize these very favorable experimental conditions. The primary difficulty stems from the fact that the point focus is not many times smaller than the distance between rings (the intensity 0.2 mm from the center of the point focus in the direction of the smaller dimension is about 0.003 times the central intensity). Hence the positions of the minima as actually obtained depend markedly on the "point focus" spot distribution and are influenced by relative intensities at the neighboring intensity peaks. Furthermore, film grain, finite range of particle diameters, and scattering due to extraneous material contribute more, percentagewise, at the minima than at the corresponding maxima. For these reasons, we have not been able to make this method of finding D (without having to assume a value for P) as reliable as the one employing the positions of the maxima and choosing P from the external consistency of the inferred values of D. The data obtained from the relative intensities at successive diffraction peaks are similarly complicated by the finite size of the point focus, but the agreement between the measured and predicted values for these relative intensities is fair. Here there is some evidence,

however, that the intensities fall off somewhat faster for large values of u than formula (1) or (2) predicts. This can be accounted for by assuming that the latex particles do not have exactly the same size, but are distributed about a mean size (as is most certainly the case).

CONCLUSIONS

K. L. Yudowitch has kindly informed us that the value of 2780A which he first obtained for the particle diameter has been revised to 2740A. This result still differs from our result of 2687.5A (based on the assumption of random orientation) by more than twenty-five times our statistical error and by about ten times the difference between our values for $P=0$ and $P=\frac{1}{2}$. We have not found any systematic errors which we believe could be of this magnitude; *it should, however, be noted that our experiment differed from that of Yudowich in that our sample was under a pressure of about 1 atmosphere of helium while his was in a vacuum.* An experiment to determine whether or not this pressure difference might account for a significant change in particle size is now being planned.§

TABLE IV. Compilation of data by individual maxima.[a]

Number of maximum	Number of measurements involved	$D_{P=0}$ (A) Uncorrected	$D_{P=0}$ (A) Corrected	$D_{P=\frac{1}{2}}$ (A) Corrected	σ_{ext} (A)	σ_{int} (A)	W
5	9	2698.4	2684.8	2699.9	5.4	5.4	343
6	11	2703.9	2696.2	2706.8	3.7	3.6	750
7	12	2692.2	2686.4	2693.9	3.4	3.3	891
8	11	2711.2	2706.7	2712.4	5.5	3.3	(see text)
9	10	2694.4	2690.9	2695.4	4.7	2.8	669
12	9	2687.2	2684.3	2687.9	2.3	3.2	1290
11	8	2690.3	2687.8	2690.8	2.2	3.2	1330
12	6	2692.5	2690.3	2692.8	2.8	3.3	1070
13	5	2690.7	2688.6	2690.7	3.2	4.1	739
14	3	2684.3	2682.2	2684.0	2.6	4.7	693
15	2	2686.6	2684.6	2686.2	1.5	5.9	540
16	2	2681.3	2679.3	2680.8	7.9	8.0	158
17	1	2686.8	2684.8	2686.0	—	13.0	59

[a] Treat the data in this table as independent, and we obtain

$\overline{D}_{P=0}=2687.5$ A, $\quad \overline{\sigma}_{ext}=1.2$ A,

$\overline{D}_{P=\frac{1}{2}}=2692.0$ A, $\quad \overline{\sigma}_{ext}=1.8$ A.

§ *Note added in proof:*—Our experiments have now shown that the change in particle diameter due to a change of 1 atmosphere in the external pressure is not more than about 3A.

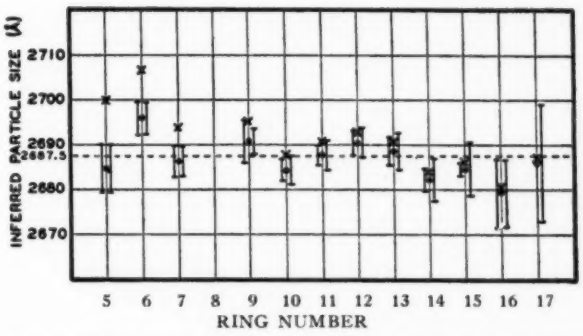


FIG. 7. This figure shows a plot of inferred particle diameters as a function of diffraction ring number as calculated from Eq. (1) both for $P=0$ (indicated by circles) and for $P=\frac{1}{2}$ (indicated by crosses). The lengths of the vertical lines on the left and right sides of the circles represent the external and internal precision indices, respectively. These indices are the same for $P=\frac{1}{2}$ as for $P=0$. Ring number 8 has been omitted as explained in the text.

OTHER WORK NOW IN PROGRESS USING THE POINT FOCUSING MONOCHROMATOR

We include with this paper a brief mention of other work now in progress with this new instrument. Preliminary results have been obtained using the well-oriented collagen fibrils of kangaroo tail tendon as scatterer, and it appears that the low angle patterns can be obtained with higher resolution, but with somewhat less intensity, than reported by other observers.[8] Measurements of the swelling properties of certain clays and of the particle sizes of some of the bacterio-phages are under consideration, and work has been started to discover the mode of aggregation of the hemoglobin molecules in the case of sickle-cell anaemia in blood corpuscles.

ACKNOWLEDGMENT

We take this opportunity to express our gratitude to the sponsors of this research (the ONR and the AEC) for their generous support and our deep appreciation to Professor Linus Pauling of this Institute, who recommended its support and encouragement to our sponsors when its success was still uncertain. We are also happy to have this opportunity to thank Professor R. B. Corey of the C.I.T. Chemistry Department and Professors J. Weigle and G. W. Beadle of the C.I.T. Biology Department for most helpful advice and discussions regarding future problems suitable for the new instrument.

[8] See, for example, R. S. Bear and O. E. A. Bolduan, J. Appl. Phys. **22**, 191 (1951).

Anelasticity of Zinc

SHUICHIRO TAKAHASHI
Taihei Mining and Metallurgical Laboratory, Omiya City, Saitama Prefecture, Japan
(Received February 18, 1952)

Anelastic properties of zinc single crystal of high purity (99.996 percent) have been studied, using bars vibrating transversely at audio frequencies in free-free modes. Internal friction, effective elasticity, and shape of resonance curve were determined against the strain amplitude. When the driving force exceeds a critical value, the internal friction increases, its maximum-amplitude-frequency decreases, and its resonance curve becomes asymmetrical. Furthermore, under this condition *a discontinuous phenomenon*, which is supposed to be closely correlated with slips occurring in the crystal, has been found. This discontinuous phenomenon was recorded under different stresses.

THE first systematic observation of internal friction of pure single crystals was due to T. A. Read;[1] his work was extended by Marx and Koehler,[2] Nowick,[3] and other investigators. They showed that there is a component of internal friction that is closely related to plastic flow. On the other hand, discontinuous jumps in creep process or tensile testing of single crystal were observed by many investigators, most notably by Becker and Orowan,[4] Schmid and Valouch,[5] and Kawada.[6]

In the present study, anelastic properties of zinc single crystal of high purity (99.996 percent) have been investigated, using bars vibrating transversely at audio frequencies in the "free-free" modes. With these specimens a discontinuous phenomenon, which is supposed to be closely correlated with slips occurring in the crystal, has been found.

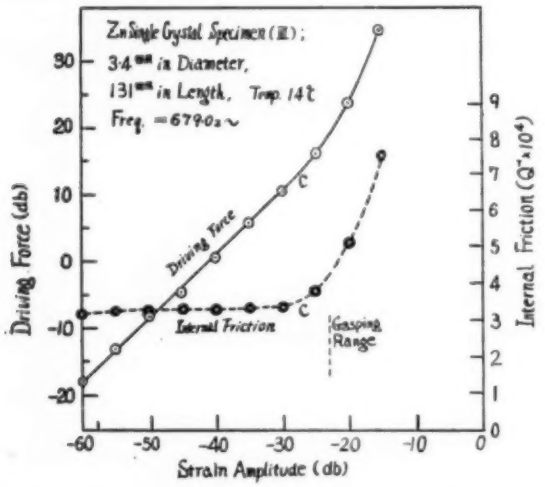


FIG. 1. Relation between driving force and strain amplitude, and dependency of internal friction upon strain amplitude for zinc single crystal.
The position of critical shearing point C varies according to the pretreatment of the specimen. The specimen used is No. 3.

The single crystal specimens were prepared utilizing the moving furnace method described by Chalmers. Its dimensions were about 3 mm in diameter and 100 mm to 250 mm in length. The specimen was suspended at the nodal lines corresponding to each mode of vibration with two fine threads. The driving voltage and the receiving voltage were led to the vertical axis and the horizontal axis of the cathode-ray oscillograph, respectively. The phase difference of the strain of the specimen and the driving force was determined from this Lissajous' figure. The strain amplitude was recorded by the electromagnetic oscillograph. The internal friction was measured by either the resonance curve method or the natural damping method; the experiments showed that they both led to the same results within experimental error. The measurements were performed in the strain range of 10^{-7} to 10^{-4} and at room temperature.

The specimen was set in vibration with increasing amplitude, where the amplitude is raised step by step. The amplitude of the specimen was adjusted so as to obtain a constant vertical amplitude on the screen of the cathode-ray oscillograph by changing the gain of the receiving amplifier. Figure 1 shows the relation between the driving force and the strain amplitude, and the dependency of the internal friction upon the strain amplitude.

(1) When the driving force, whose frequency is almost the same as the resonance frequency of the specimen, is small, the vibrating amplitude of the specimen is proportional to the driving force. The internal friction shows almost a constant value, and its resonance frequency remains constant, coinciding with the frequency at which the amplitude of the specimen is maximum. Its resonance curve is symmetrical. This aspect is shown in Fig. 2.

(2) When the driving force exceeds a critical value that seems to correspond to the critical shearing stress of the crystal, the strain amplitude of the specimen is not in proportion to the driving force; its rate of increase becomes smaller than that of the latter. With increasing strain amplitude, the internal friction increases, and the resonance curve becomes asymmetrical. These behaviors are also shown in Figs. 1 to 3. It was

[1] J. A. Read, Phys. Rev. **58**, 371 (1940); Trans. Am. Inst. Mining Met. Engrs. **143**, 30 (1914).
[2] J. Marx and J. S. Koehler, *A Symposium on the Plastic Deformation of Crystalline Solids*, Mellon Institute (1950), 171.
[3] A. S. Nowick, *A Symposium on the Plastic Deformation of Crystalline Solids*, Mellon Institute (1950), 155.
[4] R. S. Becker and E. Orowan, Z. Physik. **29**, 566 (1932).
[5] E. Schmid and M. A. Valouch, Z. Physik. **75**, 531 (1932).
[6] K. Kawada, J. Phys. Soc. Japan, **5**, 60 (1950).

found, in addition, that the resonance curves obtained with increasing and decreasing frequencies were not identical to each other, as Marx and Koehler showed on lead single crystals.[2] Figure 3 shows one of these asymmetrical resonance curves. The curve (1) in the figure is the resonance frequency curve obtained from the higher frequency side, and the curve (2) is that from the lower side. The lower frequency parts of these curves are very different from each other; their shapes depend on the speed of varying the frequency.

(3) When the driving force is somewhat stronger than the above-mentioned critical stress of the specimen, the amplitude of the specimen begins to "fluctuate" with indefinite period, although the driving force is maintained strictly constant. These "fluctuating" or "gasping" phenomena under many different stresses were recorded in various durations, sometimes up to 36 hours, sometimes down to a few seconds. Some of these "gaspings" are shown in Figs. 4(a) to 5(c). When the

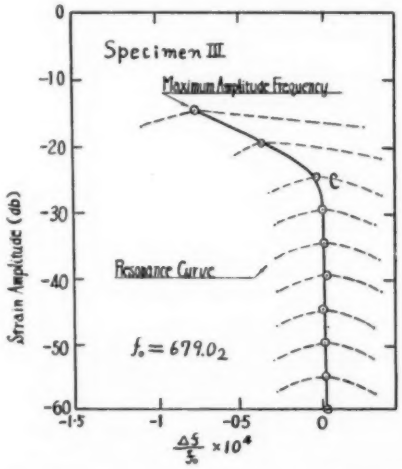


FIG. 2. Dependency of maximum strain amplitude frequency upon strain amplitude, together with resonance curves. When the driving force exceeds the critical value C, resonance curve becomes asymmetrical. The specimen is the same one as shown in Fig. 1.

driving force is only slightly stronger than the critical stress, the difference of maximum and minimum strain amplitude owing to the "gasping" is very small. Then, with increasing driving force, this difference increases to a maximum value. Speaking in general, a large "gasp" seems to be followed by a few small "gasps." The "gasping" figure is supposed to vary a little from specimen to specimen. The "gasping" range of frequency is also shown in Fig. 3. Change of Lissajous' figure mentioned previously, showed that the phase difference increases when the amplitude begins to fall and decreases when the amplitude begins to rise again. It also seems that the same change of the amplitude is not always accompanied with the same change of the phase difference. This relation seems to vary with different specimens and also with the treatments of them.

"Gasping" phenomena are also found with specimens consisting of only a few single crystals.

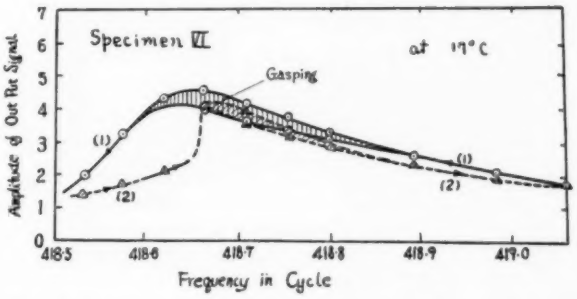
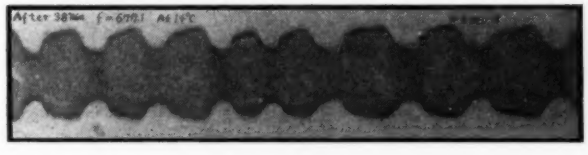


FIG. 3. Asymmetrical resonance curve for a zinc single crystal in anelastic range. The curve (1) is obtained from the higher frequency side, and the curve (2) from the lower side. The specimen used is No. 6.

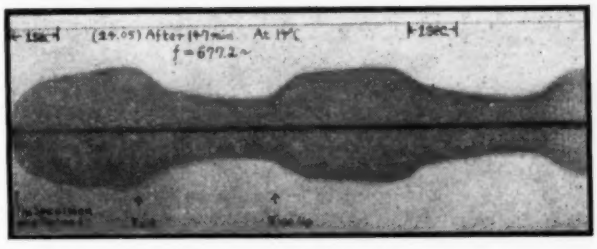
Taking account of the large value of the internal friction, the asymmetry of the resonance curve, and its irreversibility at higher stress level, it seems that the "gasping" phenomenon can be ascribed to the slips occurring in the crystal. If this interpretation is correct, a few small gasps following a large gasp show that the critical shearing stress has a lower level within some interval after large slip. It can be also expected that, in the process of slip, the Young's modulus is decreased, and the internal friction is increased. It is hoped that much information about the mechanism underlying slip process will be obtained by analyzing these dynamical behaviors.

The author is grateful to Mr. K. Asada, director of



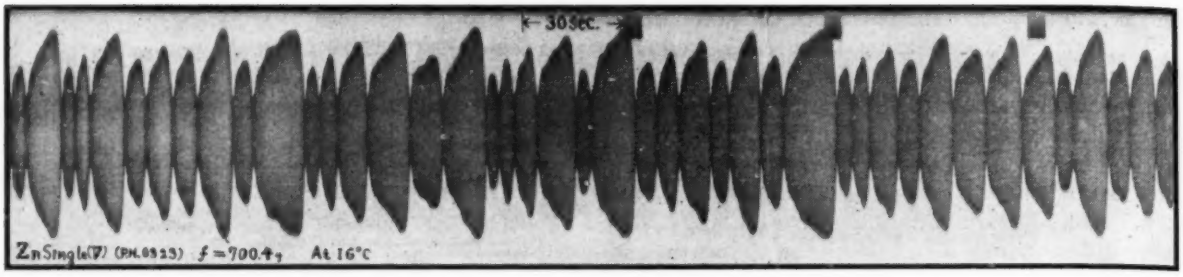
(a)

(b)

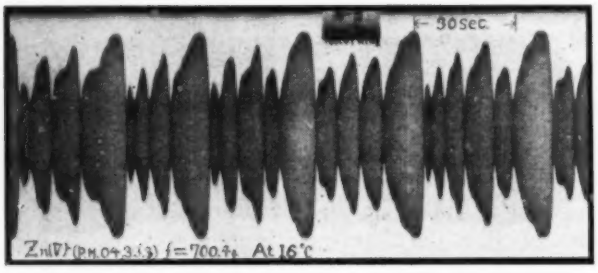


(c)

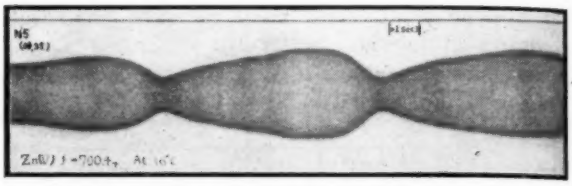
FIG. 4. "Gasping figures" of the strain amplitude of zinc single crystal under a strictly constant force in anelastic range at room temperature. The specimen used is No. 3. Figures 4 (a), (b), and (c) are the figures obtained about 38 min, 67 min, and 147 min after the driving, respectively.



(a)



(b)



(c)

FIG. 5. "Gasping" figures of the strain amplitude of zinc single crystal under a strictly constant driving force in anelastic range at room temperature. The specimen used is No. 5. Figures 5 (a) and (b) are the figures obtained about 10 min and 20.3 min after the driving, respectively. Figure 5 (c) is the figure recorded in short time at 21.8 min, which is also shown in Fig. 5 (b).

our laboratory, for his support and his permission to publish this article, and to Mr. E. Watanabe, chief of the first metal division, for his continuous interest and support. The author wishes to express his sincere thanks to Mr. A. Ōkawa, assistant professor of the Gakushuin University, for his encouragement and discussions throughout this work. The author also thanks Mr. H. Konno for his continuous help.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 8 AUGUST, 1952

The Decay and Recovery of the Pulsed Emission of Oxide-Coated Cathodes

R. M. MATHESON AND L. S. NERGAARD
Radio Corporation of America, RCA Laboratories Division, Princeton, New Jersey
(Received February 19, 1952)

The decay and recovery of the pulsed emission of cathodes consisting of BaO on nickel are studied. The emission of the cathodes is measured in diodes of cylindrical geometry having water-cooled copper anodes spaced 0.019 inch from the cathode.

Most of the diodes show a decay of anode current when first put in pulse operation. Decays of 2–50 percent are observed during 150-μs pulses. The decay tends to disappear when the tubes are aged under pulse conditions. The peak current is a linearly decreasing function of the logarithm of the duty cycle, even when there is no perceptible decay during the pulses. The rate of decrease is independent of the pulse repetition rate and decreases with cathode temperature.

Current-voltage characteristics during the pulse and recovery periods are obtained by short sampling pulses. A comparison of these characteristics with computed ideal characteristics does not discriminate unambiguously between a variation of emission and a variation of internal cathode impedance as a source of current decay. Studies to separate the contributions of these effects are in progress.

I. INTRODUCTION

IT is a common observation that the pulsed and dc emissions from an oxide cathode in an ordinary vacuum tube differ by about an order of magnitude. By the exercise of great care, Fan[1] has been able to achieve the same current from an oxide cathode under dc and pulse conditions at cathode temperatures under 1000°K, and Fineman and Dillinger have achieved dc emissions in excess of 10 amperes per square centimeter at temperatures over 1000°K.[2] However, even with careful attention to choice of materials and cleanliness, marked differences in dc and pulse currents are found. In these tubes, even if no decay is observed in microsecond pulses, it seems certain that the initial high current on application of a "unit function" voltage must decrease to the dc value. R. L. Sproull has studied the form of this decay under pulse conditions.[3] He suggests that the decay may be due to the electrolytic depletion of a barium layer at the surface of the cathode. However, other hypotheses which do not involve changes in the emitting surface lead to the same form of decay.

It is the object of the present study to locate the source or sources of decay. Because any satisfactory theory of decay must also account for the recovery of emission, a considerable portion of the effort has been devoted to the investigation of recovery rates. While the primary object has yet to be realized, a report on some of the pulse measurements may be of interest to other workers in the field. The data have also been found to be of practical value in that they provide a basis for estimates of the emission available when a cathode is subjected to groups of pulses, as in coded-pulse communications systems.

Measurements have been made of the variation of peak current with pulse length and repetition rate, of the decay of current when a "unit function" voltage is applied, and of the rate of recovery, the last by application of narrow sampling pulses subsequent to relatively long pulses.These sampling pulses have been used to measure current-voltage characteristics during decay and recovery periods in an attempt to distinguish between emission failure and an increase of internal cathode impedance. To be significant, these characteristics must be determined in a time short compared with the decay or recovery periods of the cathode. This consideration restricts the design of measuring circuits and experimental diodes.

II. TEST DIODES

The electrical design of the test diodes for emission measurements is circumscribed by three major factors: bandwidth, dissipation, and power supply regulation. The inevitable stray capacitances in the measuring circuits make it necessary to work at low impedance levels if the true voltage and current wave forms are to be observed when microsecond pulses are employed. This implies that the cathode under study should have an area large enough to deliver currents of the order of amperes. On the other hand, if the cathode area is made very large, the cost of a power supply which will deliver the required pulse current at constant voltage becomes prohibitive. When the current density from a cathode is high, as it is under pulse conditions, the instantaneous dissipation density at the anode is also high. If the duty exceeds a fraction of a percent, average dissipation becomes large and anode design a problem. This difficulty can be alleviated in part by using a close-spaced tube so that the tube voltage drop is small. Even so, the dissipation problem is such that water-cooling of the anode is necessary. The use of a close-spaced tube has another advantage when mechanical considerations lead to a cylindrical geometry, in that the diode becomes substantially planar. The theory of planar diodes is well established over the entire range of current from the Boltzmann region to the Schottky region. Hence, the experimental data

[1] H. Y. Fan, J. Appl. Phys. **14**, 552–560 (1943).
[2] Cited by A. Eisenstein in *Advances in Electronics* (Academic Press, Inc., New York, 1948), Vol. 1.
[3] R. L. Sproull, Phys. Rev. **67**, 166–178 (1945).

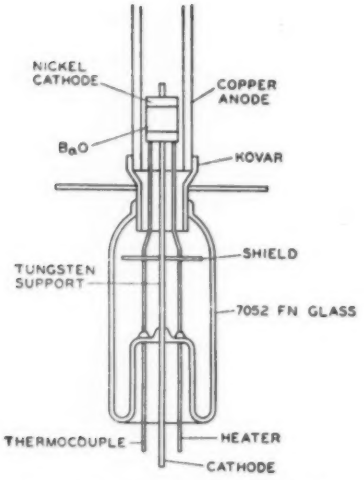


FIG. 1. Laboratory diode H-4354.

can be compared with those calculated for an ideal tube. These considerations led to the adoption of a close-spaced, cylindrical diode, with a cathode area of about 0.8 cm².

The choice of materials was largely determined by the desire to work, at least initially, with as simple a chemical system as was compatible with commercially available materials. An electrolytic nickel (RCA N-81) was chosen for the cathode base metal, and a single carbonate, $BaCO_3$ (Mallinckrodt Ultrapure) was selected for the coating. To avoid possible contamination by heater insulation, a radiation heater was used. Dissipation considerations dictated the use of a copper anode.

The test diode structure is shown in Fig. 1. The cathode is a seamless nickel cylinder 1 cm long and 0.6 cm in diameter capped by nickel thimbles. A centrally located band 0.4 cm wide is coated with 8 mg/cm² of barium carbonate. A molybdenum-nickel thermojunction is welded to the inner wall of the cathode sleeve opposite the center of the coated band. The cathode is heated by a spiral tungsten heater extending almost the full length of the sleeve. The assembly is supported by a heavy tungsten rod along the axis. This rod serves to align the cathode structure accurately within the anode during seal-in. The anode cylinder is made of OFHC copper and has an inner diameter such that the cathode-anode spacing is 0.048 cm. The anode is long enough so that the tube may be sealed off from the vacuum system by a copper pinch-off without disturbing the cathode-anode spacing.

All tubes were exhausted and baked at 300°C until the pressure fell to less than 10^{-7} mm Hg. The oven was then removed and the cathode heated to approximately 600°K and maintained at that temperature until the pressure fell below 10^{-7} mm Hg. From this point on, two procedures were used. In the first, the cathode temperature was raised abruptly to 1200°K without regard to pressure. When the pressure again fell to 10^{-7} mm Hg, the conversion of the carbonates was considered complete. In the second procedure, the cathode temperature was raised gradually from 600° to 1200°K in such a manner that the pressure never exceeded 10^{-4} mm Hg. Conversion, as indicated by evolution of gas, appeared to be substantially complete by the time a temperature of 1100°K was reached. This process required a period of three to six hours. Both methods produced equally active cathodes.

Activation was carried out by drawing gradually increasing currents. Some cathodes appeared to be fully activated on completion of the carbonate conversion and showed no departure from the 3/2 power law until high current densities were reached. Others proved difficult to activate, gave indication of poisoning when anode voltages greater than 8 or 9 volts were applied, and required prolonged heating and bombardment of the anode before activation was achieved. No satisfactory correlation of cleaning and processing procedures with activation has been possible. However, simple washing in acetone and water appeared to be as good as, or better than, extensive chemical cleaning. In both cases all parts were hydrogen-fired prior to assembly.

Cathode temperatures have not been determined with precision. To avoid complicating the tube structure, the cold junction of the thermocouple was placed inside the tube. Hence thermocouple measurements were used only to reproduce temperatures established by other means. Since the cathode surface is not visible in these tubes, optical pyrometer measurements could not be made. To establish an approximate absolute temperature scale a substantially identical structure with fine holes in the anode was wholly enclosed in a glass envelope and processed in the normal manner. Observations of cathode temperature as a function of heater power were thus obtained. To minimize errors due to anode heating the exterior of this anode was carbon coated. This, combined with the high reflectivity of the interior surface of the anode, results in small error from this source. This error is such as to make reported temperatures higher than the true temperature. The resulting curve of cathode temperature *versus* heater power agrees well with the universal curve developed by Haller.[4]

III. GENERAL OBSERVATIONS

For convenience, the term "decay" will be used in a restricted sense. When a tube operated at constant voltage shows a decrease of current during the period of application of the voltage, the tube will be said to show decay. When there is no observable decrease of current during this period, the tube will be said to show no decay, even though the pulse and dc currents may be markedly different.

Before proceeding to a detailed discussion of the measurements on decay and recovery of current in the

[4] C. E. Haller, Electronics, 126–130, 354–356 (July, 1944).

present diodes, it may be well to note briefly some general observations on the behavior of the diodes. Decay of pulsed emission has been observed with pulses ranging from 5 to 5000 microseconds. It has always been found possible to describe normal decay in terms of the Sproull formula.[3] In all cases, a reduction in cathode temperature has decreased the rate of decay and, at the same time, the peak pulse current has approached the dc value. Aging a cathode under pulse conditions at the maximum current the tube would stand without undue sparking, has caused the disappearance of decay, usually in a period of 30 to 50 hours. Similarly, aging at relatively high (0.3 A/cm²) dc currents or operating the tube as a rectifier at 60 cycles with the same average current has greatly reduced or eliminated decay. On the other hand, prolonged storage with the cathode cold, or aging for short periods on open circuit with the cathode heated, has restored or increased decay. It has been observed that the decay was always decreased immediately following a spark and that prolonged sparking, provided it was not sufficiently violent to destroy the coating, gradually reduced the decay to an undetectable level.

When the pulse length has been varied while holding the repetition rate constant, the maximum pulse current which could be drawn without sparking increased with decreasing pulse length whether or not the tube showed decay. The sparking which established the maximum permissible voltage occurred near the end of the pulse. As the maximum spark-free current was exceeded, the frequency of sparking and the duration of individual sparks increased rapidly toward the end of the pulse.

By means of the thermocouple, it has been possible to observe the cooling of the cathode by electron evaporation when a tube has operated at modest current levels. However, at high current levels, whether on pulse or dc, an increase of cathode temperature has always been observed. It is felt that the few cases of abnormal decay which have been observed may be due to heating of the cathode by pulse currents. In these abnormal cases, the tube current rose during the early part of the pulse and then declined in an apparently normal manner. Tubes showing this abnormality, which was usually transitory, were very prone to spark.

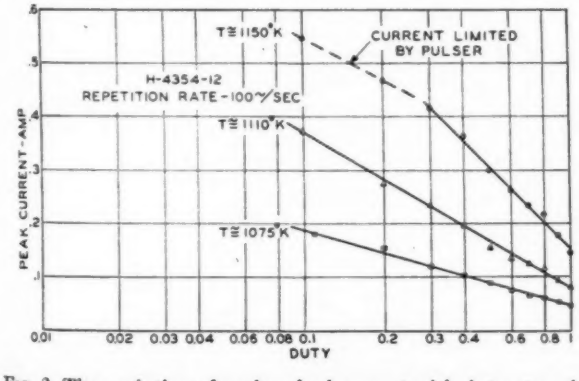


FIG. 2. The variation of peak pulsed current with duty at various temperatures with a fixed repetition rate.

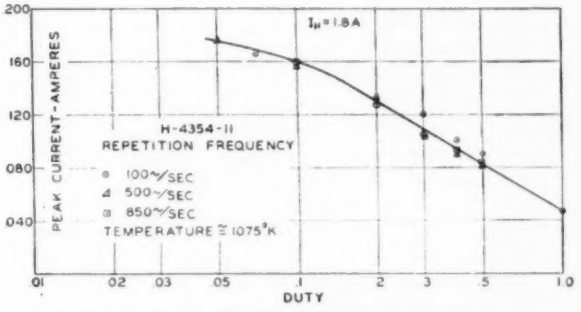


FIG. 3. The variation of peak pulsed current with duty with various repetition rates.

Prolonged aging under pulse or dc conditions has usually resulted in increased peak current. No reduction of decay, other than that caused by a reduction in temperature, has been at the expense of peak current.

IV. PEAK CURRENT VS DUTY

A possible method of obtaining data on decay and recovery of current in a diode is to observe peak current as a function of duty. This method has two major limitations. First, decay and recovery effects are not separated, so that measurements are difficult to interpret. Second, the construction of pulser and measuring circuits capable of operating over an extremely wide range of duty is a major problem. It is particularly difficult, if, as is desirable, duty is to be varied by changing both pulse width and repetition rate. A further problem is the choice of a suitable anode voltage. It must be low enough so that sparking will not occur at large duties, but at the same time it must be large enough so that the tube current is not limited by space charge alone at small duties. These conditions are frequently not compatible. For these reasons, measurements of this type have been attempted only over limited ranges of duty.

The variation of peak current with duty at constant anode voltage is shown in Fig. 2. In this case the duty has been varied by altering the pulse length while maintaining a constant repetition rate. All tubes examined under these conditions have shown the same form of variation of peak current with duty. Some of these tubes showed decays of as much as 50 percent during the pulse; in others no decay was observable. The figure shows measurements at three cathode temperatures and indicates that the rate of change of peak current with duty diminishes rapidly with decreasing temperature. The dashed section of the upper curve in Fig. 2 is not significant because the current was limited by the pulser. Changes of repetition rate over a limited range do not affect the shape or position of these curves. This is brought out in Fig. 3 which shows the variation of peak current with duty from 0.05 to unity at three repetition rates.

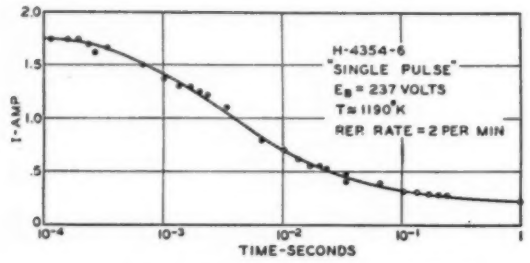


FIG. 4. The decay of anode current on the application of a constant voltage to the anode.

Plots of peak current *versus* the logarithm of the duty show an approach to zero slope for duties less than 10 percent. For sufficiently small duties, the slope must become zero, but present measurements at lower duties have not located the knee of the curve.

A few measurements of "isolated" pulses have been made to determine whether the peak currents at very low duties are indicative of the peak current which may be drawn by application of a "unit function" voltage. The decay of current with time in an "isolated" pulse is shown in Fig. 4. When the interval between pulses was varied from one-half minute to several minutes, the peak current and the rate of decay of current showed no variation. The peak current observed under these conditions was the same as that for a 5-microsecond pulse applied at a repetition rate of 100 per second (0.05 percent duty).

There are several possible mechanisms which may account for the observed decay and the variation of peak current with duty. Among these are poisoning due to the anode and poisoning due to residual gas, both of which are surface phenomena. Another possibility is the electrolytic transport of barium from the surface as suggested by Sproull. This is, in part, a volume effect but may be observed in a diode as a limitation of the zero-field emission. Further possibilities are changes in coating or interface resistance.

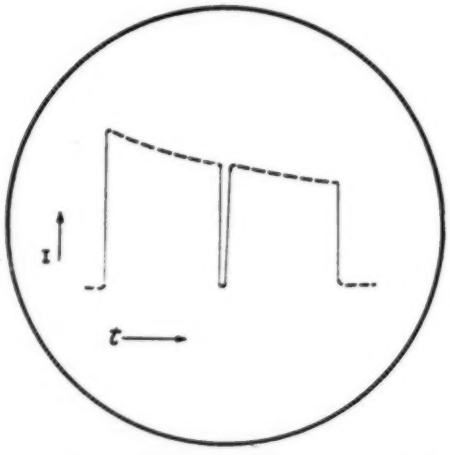


FIG. 5. An oscilloscope trace showing the decay of anode current during a 145-microsecond pulse with a 3-microsecond interrupting pulse for the measurement of an *E-I* characteristic. The timing marks along the pulse occur at 10-microsecond intervals. Note the small current recovery during the interrupting pulse.

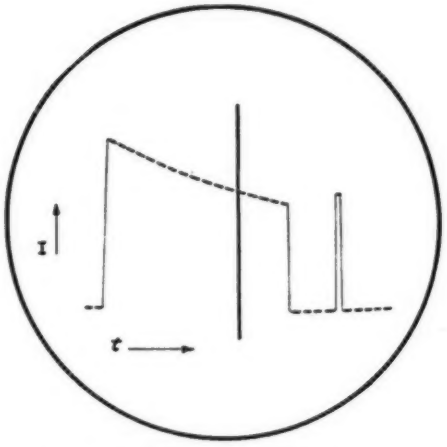


FIG. 6. An oscillograph trace showing the current decay during a 145-microsecond pulse and a 3-microsecond sampling pulse after the main pulse to measure the recovery rate of the cathode.

It would be most desirable to distinguish clearly between decays due to zero-field-emission limitation and decays due to changes of coating or interface impedance. If *E–I* characteristics can be obtained over a sufficiently wide voltage range throughout the decay and recovery periods, it should be possible to make this distinction.

V. *E-I* CHARACTERISTICS DURING DECAY AND RECOVERY

If *E–I* characteristics are to be significant in interpreting effects during decay and recovery, they must be obtained by measurements which do not perturb the decay and recovery appreciably. In the present diodes, such measurements have been possible, using pulse lengths of 100 to 300 microseconds and repetition rates of the order of 100 cps. To obtain the *E–I* characteristics during the primary pulse, this pulse was inter-

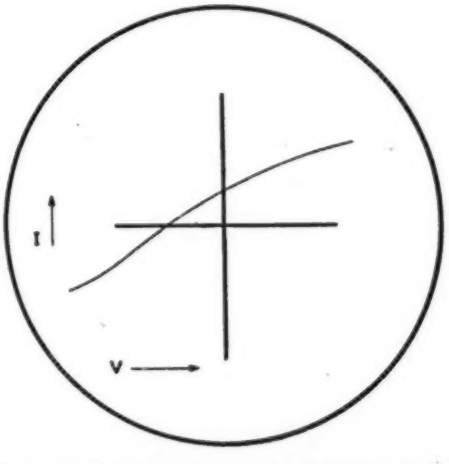


FIG. 7. An oscillograph trace of the *E-I* characteristic of the test diode as measured by the 3-microsecond sampling pulse. The heavy vertical and horizontal lines are calibrating markers.

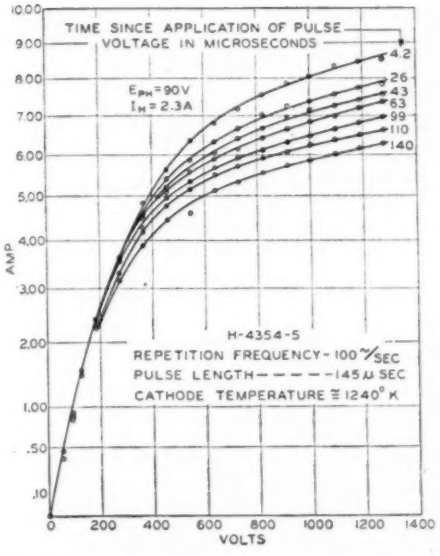


FIG. 8. *E-I* characteristics during current decay. The curves are plotted on 3/2 law paper.

rupted for approximately three microseconds. The interrupting pulse was approximately triangular, the leading edge being very steep. The *E–I* characteristics were obtained during the period of reapplication of anode voltage. Figure 5 shows the slight effect of such an interrupting pulse on the decay of current in one of the diodes.

E–I characteristics during the recovery period were obtained by application of a three-microsecond sampling pulse. In this case also, the *E–I* curve was obtained during the period of increasing anode voltage. Figure 6 shows the main pulse with the sampling pulse to the right. The height of this pulse is indicative of the extent of recovery of the diode. Figure 7 shows a typical *E-I* characteristic obtained by this method. The heavy

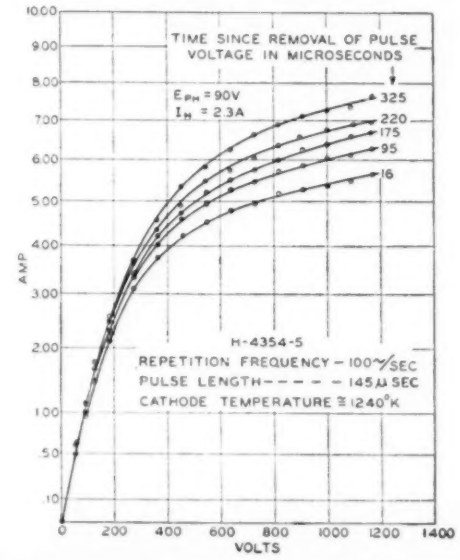


FIG. 9. *E-I* characteristics during current recovery. The curves are plotted on 3/2 law paper.

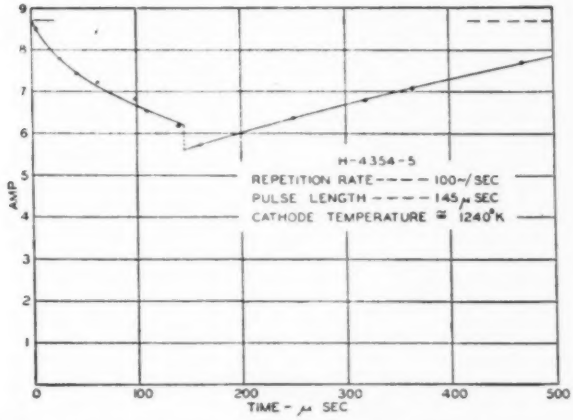


FIG. 10. The decay and recovery of the anode current of diode H-4354-5. Plotted points are the 1200-volt points of the *E-I* characteristics of Figs. 8 and 9.

horizontal and vertical lines are calibrating marks. Figure 8 is a set of *E–I* characteristics showing the change occurring during the period of the pulse. Figure 9 is a similar set of characteristics showing the change during the early part of the recovery period. In both of these sets the maximum anode voltage was adjusted to a point just less than that at which sparking near the trailing edge of the pulse occurred. The decay and recovery curve obtained from the *E–I* characteristics is shown in Fig. 10. Figure 11 shows the decay and recovery cycle of another tube. The initial recovery is more rapid in the second tube. It will be noted that in both tubes there appears to be a discontinuity in the available current after the cessation of the main pulse. An examination of the main- and sampling-pulse traces in photographs, such as Fig. 6, shows that recovery starts immediately after the main pulse and that the apparent discontinuity in the plots of Figs. 10 and 11 is a result of the slow upward drift of the current (even after the tubes had been aged under full pulse conditions for 20 hours prior to the measurements) and the order in which the readings were taken. The first recovery point was read first and the last

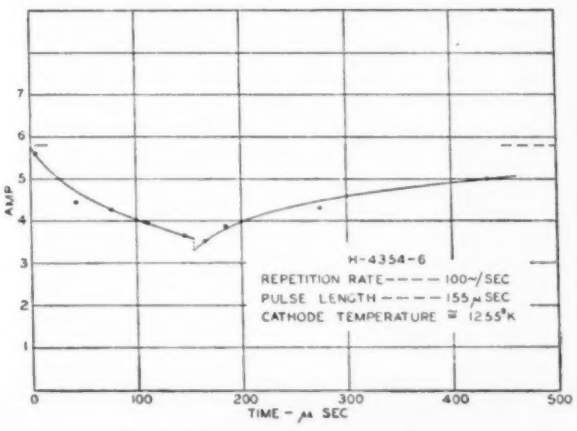


FIG. 11. The decay and recovery of the anode current of diode H-4354-6.

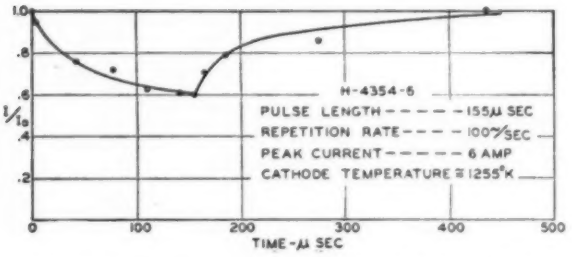


FIG. 12. Decay and recovery currents of diode H-4354-6 normalized to the initial pulse current.

decay point was read last. Hence, the discrepancy between the decay and recovery curves is the total drift. Figure 12 shows the data of Fig. 11 with the recovery points multiplied by the factor which matches the decay and recovery curves. This is not a proper adjustment because the points are not properly weighed according to elapsed time, but it gives a reasonably accurate picture of a decay and recovery cycle.

Figure 13 shows computed ideal planar diode characteristics. Below the field-free emission point ($I/I_0=1$) the characteristic satisfies the 3/2 power law. It is assumed that the voltage V_0 required to draw the field-free emission under space-charge limited conditions is large compared to KT/e so that initial velocities may be neglected in the current decade below the field-free emission current. Above the field-free current point, the characteristics are Schottky curves corrected for the effects of space charge. The effect of space charge on the field strength at the cathode surface is computed by using the zero-order solution for the general behavior of vacuum tubes as given by Llewellyn,[5] and solving for the field strength at the cathode surface as a function of current and applied voltage with the diode spacing as a parameter.[6]

The substitution of this field strength in the Schottky formula and the subsequent solution for the current in terms of the applied voltage and the tube parameters yields a result which is most conveniently expressed in terms of the normalized curves of Fig. 13. I/I_0 is the normalized current, i.e., the ratio of the actual current to the field-free emission current. V/V_0 is the normalized voltage, i.e., the ratio of the applied voltage to voltage required to obtain the field-free emission under space-charge limited conditions. γ is the parameter which embodies the remaining variables, namely, the cathode temperature and cathode-anode spacing.

In principle, it should be possible to superpose an experimental curve, plotted on logarithmic paper, on the theoretical curves so that the low current end of the experimental curve is asymptotic to the 3/2 law part of the theoretical curves and the high current end is asymptotic to a theoretical curve in the Schottky region. If such a fit were obtained, the field-free emission could be read from the theoretical curves and the discrepancy between the theoretical and experimental curves could be used to determine cathode impedance. Sparking limits the anode voltage to a value such that the data obtained do not extend far enough into the Schottky region to permit satisfactory comparison with the theoretical curves. The best fits obtainable with the present data indicate that the temperature of the cathode falls during the pulse. The rises of temperature observed with the cathode thermocouples make a temperature fall during a pulse seem most improbable. The observation that the cathode temperature does rise and the departure of the experimental curves from the 3/2 law at very modest current densities makes it quite clear that oxide or interface resistance is playing a substantial role. If it is assumed that the deviations from the space-charge line are due to a nonlinear resistance in coating or interface, the power density at the sparking point agrees quite well with values obtained in the same way by Eisenstein and others.[7] It appears that sparking, a consequence of cathode resistance, makes the determination of the cathode resistance impractical by the curve matching method.

All of the curves of Figs. 8 and 9 appear to have the same shape. This is borne out in Fig. 14 in which the curves of Figs. 8 and 9 have been normalized to the peak current at the beginning of the pulse. The spread in points is indicated at 200-volt intervals. In the high voltage region where departures from the $\frac{3}{2}$ law are greatest and where the curves might be ex-

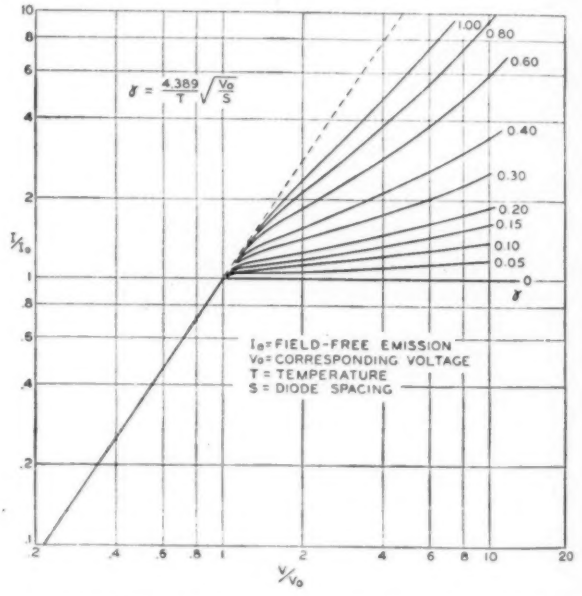


FIG. 13. Theoretical *E-I* characteristics for a planar diode with the Schottky characteristics corrected for the effect of space charge on the field at the cathode.

[5] F. B. Llewellyn, Proc. Inst. Radio Engrs. **21**, 1532–1573 (1933).

[6] A similar computation of the field at the cathode has been made by H. F. Ivey, Phys. Rev. **76**, 554 (1949).

[7] A. Eisenstein, J. Appl. Phys. **20**, 776–790 (1949).

pected to differ most, the curves become identical within the experimental accuracy. It should be noted that this region also has the highest experimental accuracy. That this behavior is not characteristic of simple emission decay can be seen from the following argument: As the field-free emission of a cathode falls, V_0 of Fig. 13 falls as the $\frac{2}{3}$ power of the emission current, γ decreases as the $\frac{1}{3}$ power of the emission current, and the characteristic flattens towards the voltage axis. This flattening is not observed in the experimental curves. While Fig. 14 sheds no new light on the mechanism responsible for decay, it supports the previous conclusion that the interior of the cathode plays a significant role in the decay process.

CONCLUSIONS

Although the principal object, to locate definitely the source or sources of decay, has not been achieved, a large body of pertinent data has been accumulated. Any satisfactory theory of decay must explain these data.

While the present data do not definitely locate the source of the decay, they do indicate that some possible sources are improbable. If the anode is assumed to be the source of decay, then it is difficult to explain the fact that sparking observed here and elsewhere can be correlated with definite power densities in the cathode despite widely varying tube geometries. If gas poisoning is assumed to be the source, a suitable mechanism must be devised. If any likely collision cross section for ionization is chosen for the residual gas in the cathode-anode space, then a simple calculation shows that, for any reasonable pressure, all the gas in this space will arrive at the cathode in a very small fraction of the time involved in the observed decays. If decay is assumed to be due to a change in the field-free emission of the cathode, then it would be expected that sparking could be correlated with the field strength at either the cathode or the anode. Such a correlation has not been observed in these experiments. Furthermore, the fact that the shape of the *E-I* characteristics remains substantially invariant during decay is difficult to reconcile with an emission limitation. Other possible sources of decay are coating and interface resistances. The present investigation has produced no data which are inconsistent with the assumption that decay and sparking are caused by internal cathode impedance.

Further measurements are planned to distinguish between changes of internal cathode impedance and other sources of decay. Several approaches are being investigated. (1) Measurements of anode power dissipation and cathode temperature rise would, in theory, make the necessary distinction. The energy input

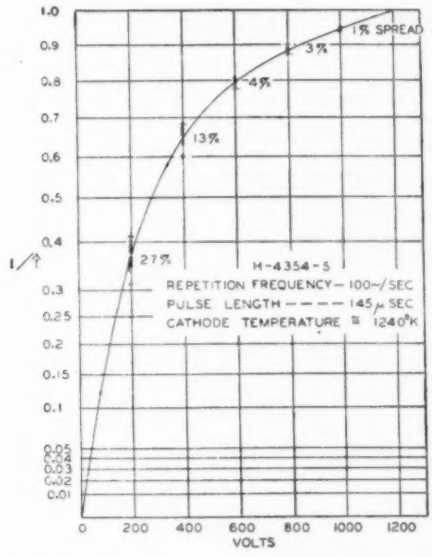


FIG. 14. Decay and recovery *E-I* characteristics normalized to the initial current.

to a diode during a pulse is the product of the charge content of the current pulse and the applied voltage. If the voltage drop through the cathode is trivial, all of the energy will be dissipated at the anode. On the other hand, if the voltage drop through the cathode is substantial, part of the energy input is dissipated in the cathode, the cathode temperature rises and the anode dissipation decreases. A measurement of cathode temperature rise or a measurement of anode dissipation, as a function of pulse length, would separate emission decay from a change of cathode resistance. However, the practical difficulties involved, such as measuring a small change in the already large anode dissipation, make this method unattractive. (2) Measurements of the surface potential of the cathode throughout the pulse, although difficult, appear the most promising. Possible methods of determining this potential are (a) by passing a beam of charged particles through the cathode-anode space parallel to the surface of the cathode and observing the change in their velocity; (b) by measurement in a retarding field of the velocity of electrons passing through a hole in the anode; and (c) by measurement of the potential of an emitting probe placed close to the cathode surface.

It is essential to complete any of these measurements in a time small compared with the period of the decay in order to establish that the decay is entirely due to changes within the cathode. This requirement complicates the measurement problem. None of the methods so far considered is amenable to detailed analysis, and further experiments will probably be necessary to select a suitable method.

Measurement of the Internal Energy in Copper Introduced by Cold Work

BENJAMIN WELBER
Lewis Laboratory, National Advisory Committee for Aeronautics, Cleveland, Ohio
(Received March 6, 1952)

A determination was made of the ratio of the stored energy to the energy expended in cold-working copper by twisting. Ratios of about 0.03 were found in three cylindrical samples which were cold-worked to the extent of 8 to 12 cal/g.

I. INTRODUCTION

WHEN a metal is cold-worked by stretching, twisting, or rolling, etc., a great number of its physical properties are altered.[1–4] The theoretical interpretation of these effects is usually based on the idea that the process of cold working changes the state of the solid by increasing the number of a special type of lattice imperfection, the dislocations.[5–11] One way of obtaining information regarding the state of the cold-worked material is the determination of the stored energy as a function of the amount of mechanical work expended in the cold-working process. If the amount of the mechanical work is known, it is possible to obtain this information either by measuring the heat released during cold working or by determining the stored energy released upon subsequent heating. This quantity can be determined with relatively good accuracy in the case of copper because its high thermal conductivity assures a more uniform temperature throughout the sample and its nearly constant specific heat simplifies the interpretation of the results. Table I summarizes the values obtained for copper in previous investigations of this sort. W represents the mechanical work per gram, averaged over the sample, which is expended during the cold-working process, and E the average energy per gram which is stored. It is these quantities, rather than the distributions of local energy densities w and e, which are obtained in experiments in which the plastic deformation is non-uniform.

TABLE I. Ratio of the stored energy to the total mechanical energy expended in cold-working copper. E, stored energy; W, mechanical energy.

	Measured quantities	W in cal/g	$\frac{E}{W}$
Farren and Taylor[a] (1925)	W, $W-E$	0.82	0.09
Sato[b] (1931)	W, E	4.2	0.042
		9.7	0.031
		18.0	0.018
Taylor and Quinney[c] (1934)	W, $W-E$	0.96	0.075
		6.2	0.086
		14.1	0.081
Quinney and Taylor[d] (1937)	W, E	5.5	0.09

[a] See reference 12.
[b] See reference 14.
[c] See reference 13.
[d] See reference 15.

The measurements reported by Farren and Taylor[12] and by Taylor and Quinney[13] were done by means of the first of the above mentioned methods. In the case of Farren and Taylor, the cold work was achieved by stretching a rod of metal by a direct load. With such a system only a very limited amount of work can be done on the specimen before it breaks, hence, the rather low value of W which is shown in the table. The temperature rise of the specimen on applying the load allowed one to estimate the latent energy retained by the specimen.

Taylor and Quinney similarly made indirect measurements of the stored energy ratio by determining the temperature rise of cylindrical rods subjected to rapid incremental twists. By this method it is possible to impart much more cold work to the sample and it is to be expected that the stored energy will eventually approach some saturation value as the amount of applied work is increased, so that the ratio of the stored to the applied energies will decrease monotonically.

Sato,[14] and also Quinney and Taylor, in a later publication,[15] employed annealing techniques for determining the amount of energy stored in twisted rods. The results reported by Sato are based on a sensitive differential method, in which two specimens of identical dimensions, one cold-worked and the other annealed, were inserted in two equal holes in a block of silver, which was then heated at a uniform rate. The specimens were insulated from the silver block by a silica lining, and on the assumption that the thermal resistance between each sample and the silver block were identical,

[1] J. W. Rutter and J. Reekie, Phys. Rev. **78**, 70 (1950).
[2] G. Borelius, Ann. Physik **60**, 381 (1919).
[3] G. Tamman and C. Wilson, Z. anorg. u. allgem. Chem. **173**, 156 (1928).
[4] B. Wiener and G. Groetzinger. J. Appl. Phys. **23**, 441 (1952).
[5] G. I. Taylor, Proc. Roy. Soc. (London) **A145**, 362 (1934).
[6] E. Orowan, Z. Physik. **89**, 634 (1934).
[7] J. S. Koehler, Phys. Rev. **75**, 106 (1949).
[8] J. K. Mackenzie and E. H. Sondheimer, Phys. Rev. **77**, 264 (1950).
[9] R. Landauer, Phys. Rev. **82**, 520 (1951).
[10] J. D. Eshelby, Proc. Roy. Soc. (London) **A197**, 396 (1949).
[11] F. R. N. Nabarro, Proc. Roy. Soc. (London) **A209**, 278 (1951).
[12] W. S. Farren and G. I. Taylor, Proc. Roy. Soc. (London) **A107**, 422 (1925).
[13] G. I. Taylor and H. Quinney, Proc. Roy. Soc. (London) **A143**, 307 (1934).
[14] S. Sato, Sci. Rep. Tohaku Imp. Univ. **20**, 140 (1931).
[15] H. Quinney and G. I. Taylor, Proc. Roy. Soc. (London) **A163**, 157 (1937).

the latent energy would be proportional to $\int(\Delta T)dt$ where ΔT is the temperature difference between the two specimens at time t. Due to the impossibility of matching exactly the two thermal resistances, corrections had to be applied which made the quantitative results somewhat doubtful. To obviate this difficulty, a direct calorimetric method was later employed by Quinney and Taylor[15] in which the sample was heated by an internal heater and the thermal loss reduced to a minimum. The temperature of the cold-worked sample was measured as a function of the amount of heat supplied to it, and the relation between temperature and energy input was compared with that resulting in the case of an unworked, or an annealed sample of the same dimensions.

The measurements of Taylor and his co-workers show roughly the same value of E/W of about 0.08, thus displaying consistency with each other and indicating that saturation effects do not affect this ratio up to the highest value of W (14 cal/g) involved in these experiments. However, the published values of $\Delta E/\Delta W$[14] which are much more indicative of saturation effects than the values E/W show that the amount of the energy stored increases more slowly with increasing W for values of W greater than 13 cal/g.

In Sato's experiment the value of E/W was in every case less than half that found in the above experiments and the tendency towards saturation seems to appear already at $W=4$ cal/g. In order to resolve this serious discrepancy it was decided to make a new determination of the stored energy in copper as a function of the mechanical energy expended. To obtain the highest possible accuracy, these measurements were carried out at rather large values of W, the limitation being fixed by the strain at which the sample fractured. It is reasonable to assume that if E/W is constant in this range it will be constant for all lower values of W. If this is the case, e/w will likewise be constant and equal to E/W, regardless of the distribution of e throughout the sample. The method of Quinney and Taylor was chosen for this investigation because it represents, in our opinion, the best and the most direct approach.

II. EXPERIMENTAL PROCEDURE

(a) Cold Working of Samples

A long cylindrical copper rod, approximately $\frac{5}{8}$ inch in diameter, was first annealed in vacuum at a temperature of 700°C for about one hour. It was then twisted in a Tinius Olsen torsion testing machine on which readings of the acting torque L and the angle of twisting ϕ at any instant could be taken (at a rate of half a revolution per minute); W was determined by integrating the measured value of L over ϕ. After the twisting, the rod was cut into sections approximately three inches in length and, in order to accommodate the heater, small $\frac{1}{8}$-in. holes extending almost all the way through the sample were drilled along the center of each. Drilling and cutting off were done at very slow speeds. Finally the sample was thoroughly cleaned and weighed, after which it was ready to be installed in the annealing apparatus.

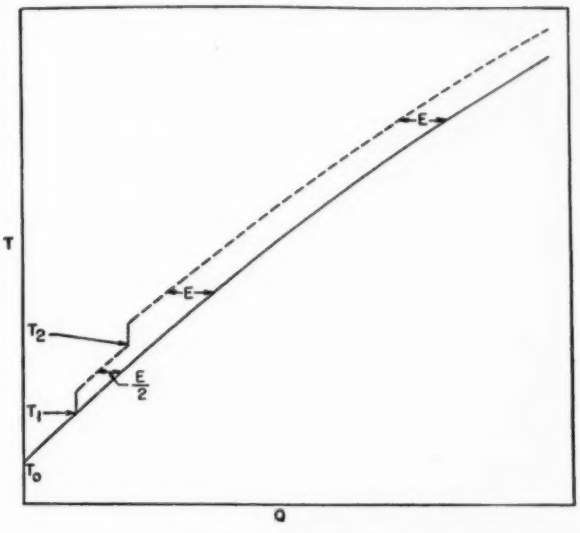


FIG. 1. Idealized T *versus* Q curves showing the two-step release of latent energy.

The mass of material which was drilled out amounted to about four percent of the total mass. The presence of this material during the twisting will make little difference in the mechanical work absorbed by the sample, since the strains produced in the regions near the axis are small compared to those in the rest of the sample. The work expended in this part of the sample is, therefore, being taken as zero.

(b) General Remarks Concerning Annealing Experiments

If an amount of heat Q is supplied per gram of the sample, the temperature rise will be a linear function of Q as long as the specific heat c is independent of the temperature and the heat losses are negligible. With copper the first condition is practically fulfilled for the temperature range in question.

The effect of heat losses may be demonstrated by consideration of an idealized experiment in which a cold-worked sample is to be compared with an otherwise identical annealed sample in order to determine the stored energy E in the cold-worked sample. We consider first the case of the annealed sample which is heated at a constant rate of energy input dQ/dt, starting at room temperature. If the heat losses are proportional to the difference between the temperature of the sample and room temperature, then the temperature rise is that shown by the full curve of Fig. 1 and is given by the equation

$$(T-T_0)=T_{\max}(1-e^{-aQ}), \qquad (1)$$

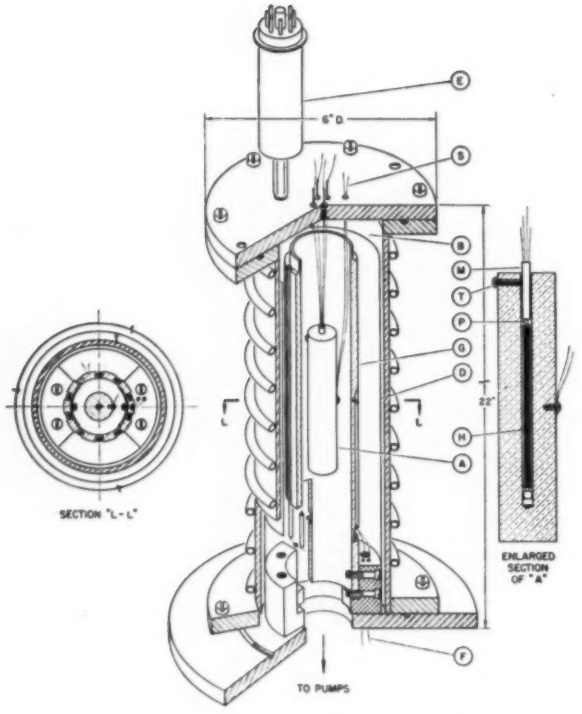


FIG. 2. Annealing apparatus and sample.

where a is a constant. The initial slope of the curve represents $1/c$.

The cold-worked sample is then heated in the same fashion. If we assume (for the sake of simplicity) that the same amount $E/2$ of latent heat is released instantaneously by this latter sample at each of the two temperatures T_1 and T_2, then the temperature of the sample is given by the solid curve of Fig. 1 at temperatures below T_1, and by the dotted curve above this temperature. If T_1 and T_2 are in the nearly linear section of the curve, the separation of points on the two curves corresponding to the same temperatures, $(Q_{\text{ann}}-Q_{cw})$, is equal to $E/2$ below and to E above the temperature T_2, respectively. It should be noted that this separation remains constant even in the nonlinear part of the curves although it becomes more and more difficult to determine accurately as the slope decreases. The above results can be easily generalized for the case of a continuous release of latent energy.

In an experiment designed to determine the total amount of the stored energy, it is therefore essential to make the heat losses so small that the temperature at which the latent energy is completely released is still in the linear part of the $T-Q$ curve. For the case of the specific heat varying with temperature, an indication of low heat loss is that the slope of the annealed specimen be close to $1/c$.

If the heat losses are considered to be unchanged for successive runs, then the error can be determined by considering the following expression for the stored energy:

$$E=Q_{\text{ann}}(T', T'')-Q_{cw}(T', T''), \tag{2}$$

where $Q(T', T'')$ is the energy per gram required to bring the temperature of the sample from T' to T''. T' is a temperature lower than the one at which stored energy begins to be released, and T'' a temperature higher than the temperature at which the release is completed. It follows that if the inherent random error in measuring the quantities $Q(T', T'')$ with T' and T'' exactly known is ΔQ, and if the error in measuring T' (or T'') is ΔT, then the error ΔE associated with a given measurement of E is

$$\Delta E=(2\Delta Q^2+4c^2\Delta T^2)^{\frac{1}{2}}. \tag{3}$$

Since from a single run one obtains values of $Q(T', T'')$ for a number of different values of T' and T'', the E determined in a single experiment consisting of two runs is in error by considerably less than is indicated by Eq. (3). The error in the measurement of the temperature arises chiefly from the change in the characteristics of the thermocouple in the course of a run. If we ignore all other contributions to ΔT, the measurements of thermocouple stability made by Dahl[16] imply a value no greater than 0.1°C for this quantity. This amounts to an error of no more than 4 percent in a single measurement of E in copper if E is approximately 0.5 cal/g. It is more difficult to estimate the quantity ΔQ, but a simple calculation shows that if the measurement of Q is accurate to one part in fifteen hundred, a value of E of the order of 0.5 cal/g obtained in a single measurement will be in error by no more than 10 percent from this cause.

(c) Arrangement of Annealing Apparatus

The schematic diagram in Fig. 2, which is compressed in height relative to the horizontal dimensions as indicated on the drawing, shows the experimental set-up. A represents the sample in the form of a cylindrical rod about three inches long, which is supported in the vacuum chamber B solely by the heater leads in order to reduce conduction losses.

Energy to the sample is supplied by a heater H which is contained within a small hole in the sample. The heater and the current supply leads consist of 34 inches of number 28 nichrome wire. In order to determine the energy used to heat the sample, two "voltage" leads are attached at points P. The heater is nowhere in direct contact with the sample. Its lower end rests on a ceramic bead inside the sample without touching it otherwise. The ceramic tube M which separates the four heater wires fits snugly into the sample hole and is held by a set screw. In order to reduce the heat losses, the sample is surrounded by a "guard-ring" furnace, G, which rests on the pair of steatite blocks.

[16] A. I. Dahl, *Temperature, Its Measurement and Control*, (American Institute of Physics, Reinhold Publishing Corporation, New York, 1941), p. 1238.

It was constructed of a 20-inch length $1\frac{1}{2}$-in. copper pipe which was highly polished on the inside and has 20 linear elements consisting of number 22 nickel wire encased in flat ceramic tubes lying in milled slots on its outer surface. The efficiency of the radiation shielding thus achieved is indicated by the fact that the solid angle formed by each open end of the furnace as viewed from the sample is of the order of $10^{-2}\pi$ steradians. The temperature of the furnace is measured by the thermocouple F. A thin, highly polished nickel cylinder D is inserted in the vacuum container to reduce the radiation loss from the furnace. The electrical leads for the sample heater, the furnace, and the thermocouples are brought out at the top and the base of the vacuum chamber through suitable Kovar steels.

The chamber and flanges are of stainless steel, and are water-cooled on the outside by 17 feet of $\frac{1}{4}$-in. copper tubing. A D.P.I. type VMF-20 oil diffusion pump backed by a roughing pump was sufficient to keep the pressure below the readable range of the thermocouple gauge E, even during the first operation of a new heater or furnace.

The wires for the two chromel-alumel thermocouples were cut from the same spools (number 28 wire). The couples were calibrated only once, and left undisturbed thereafter. The temperature difference between the hot junctions of S and F was directly monitored with a Rubicon box galvanometer (resistance—15 ohms, voltage sensitivity, 2.0 μV/mm). (See Fig. 3.) The two sets of cold junctions were maintained at 0°C by two separate ice baths.

(d) Performance of the Annealing Experiment

To measure the energy supplied, it was necessary to know the value of $\int_0^{t(T)} VI dt$ for the elasped time t during which the temperature of the sample rises from the initial temperature T_0 to the temperature T. The determination of this quantity is greatly simplified if the product VI is kept independent of the time. Due to the change of the heater resistance R with temperature, this product cannot be kept constant either by using a constant voltage or a constant current source.* In the present experiment, an arrangement was used which kept the current nearly constant during extended periods so that only the voltage had to be recorded at frequent intervals (once every two minutes). The diagram of the circuit is shown in Fig. 3. The voltage source consisted of a set of lead storage batteries of 200 ampere-hour capacity supplying 42 volts. The ballast resistance consisted of five Westinghouse type 788 ballast tubes in parallel, which kept the current of approximately one ampere constant to one part in three thousand in the course of an experiment. The voltage and the current were measured with a Rubicon potentiometer by use of a voltage divider and a 0.1 ohm standard resistance, respectively. A few hours before the start of an experiment the dummy load shown in Fig. 3 was connected in place of the heater, so that by the beginning of the experiment, the power supply had achieved its stable operating condition.

It was found most practical to heat the sample at a rate of about 10°C per minute. The maximum temperature attained was between 450° and 500°C so that the duration of the experiment was about 45 minutes. Two operators were required, the one to maintain the temperature of the "guard ring" furnace close to that of the sample by manual adjustment of a variac, the other to make the necessary measurements of sample temperature, and heater voltage and current. With some previous practice, it was possible, throughout a run, to keep the "guard ring" furnace within ±0.25°C of the temperature of the sample.

A small fraction of the heat supplied is used to raise the temperature of the heater which necessitates an appropriate correction. The total weight of the heater and the ceramic bushing M is approximately one percent of the weight of a sample. The specific heat of the ceramic is, however, two to three times that of copper so that the heat capacity of the heater is of the order of two percent of that of the sample. Accordingly, the quantity Q, the heat supplied per gram of the sample, was taken as 98 percent of the total heat supplied to the heater per gram of sample.

III. RESULTS

Figure 4 contains the results of a typical measurement on the sample which had been cold worked to the extent of 9.4 cal/g. The figure shows the relation between the temperature T and the heat supplied per

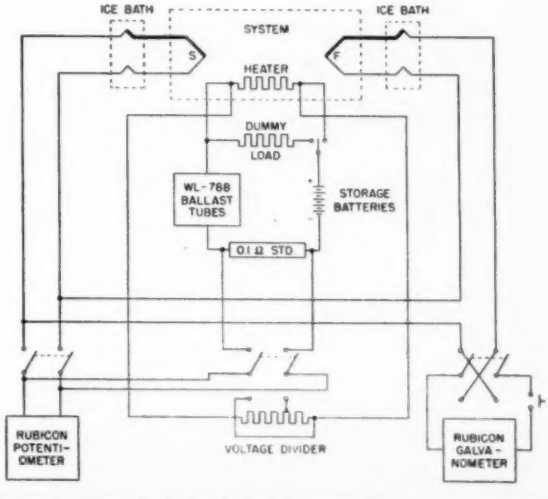


FIG. 3. Electrical circuit for the annealing measurements.

* A simple way to obtain approximately the condition of constant power input for a limited change of the heater resistance is by the use of a constant resistance equal to the initial value R_0 of R in series with the heater and a constant voltage source. Starting with this condition, an increase of the resistance $\Delta R/R$ of 0.1 e.g., decreases the power by a factor of 0.0025 only. Since these conditions cannot be relied upon to exist throughout the duration of an experiment, it is still necessary to make accurate voltage and current measurement, so that little is gained.

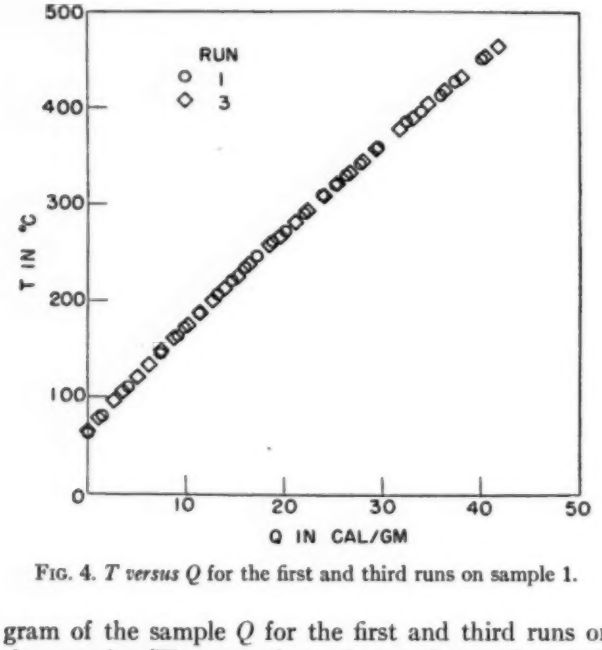


FIG. 4. *T versus Q* for the first and third runs on sample 1.

gram of the sample Q for the first and third runs on the sample. (The second run is not shown since it is practically identical with the third, and the scale of the drawing makes the inclusion of the extra points inconvenient.) In the first run, the annealing of the originally cold-worked sample occurs. 50 to 60°C is used as initial temperature for this and succeeding curves since the amount of energy supplied to the heater in the first few minutes of an annealing run is difficult to measure accurately on account of the rapid change of its electric resistance. Furthermore, after this temperature has been reached, transient effects in the temperature distribution, which occur after the heater is energized, do not make themselves felt any more. This follows from the near linearity of the curves for the sample in its annealed state.

From the slope of the T *vs* Q curves for the annealed samples, the specific heat for copper can be computed as a function of temperature. The specific heats obtained in this way are bound to be too high on account of the heat losses of the sample by conduction through the suspension wires, etc., and therefore, this deviation from the correct value provides an estimate of such losses. The average specific heats uncorrected for heat losses obtained from all the experiments are, therefore, listed in Table 2 for various temperatures together with the accepted values.[17] The agreement between the measured values and the accepted values is sufficiently good up to the highest temperature at which latent energy is released as will be seen below. This means that the heat losses were small enough not to affect the determination of E.

TABLE II. Accepted values C of the specific heat of copper and apparent values C_a (uncorrected for heat losses) following from the annealing experiments, as a function of temperature.

T (°C)	C_a (cal/g°C)	C (cal/g°C)
100	0.097	0.1019
200	0.101	0.1020
250	0.109	0.1023
300	0.108	0.1026
400	0.113	0.1032

In Fig. 5 the quantities $H(T)=(Q_{\text{ann}}''-Q_{cw})_T$ and $(Q_{\text{ann}}''-Q_{\text{ann}}')_T$ are plotted for three samples whereby the quantities Q_{ann}' and Q_{ann}'' refer to the two runs made with the sample in an annealed state (second and third run). If the sample is completely annealed after the end of the first run, the second and third run should give identical results, and the quantity $(Q_{\text{ann}}''-Q_{\text{ann}}')_T$ provides essentially a measure for the errors involved in the annealing experiments. The experimental curves shown indicate that the annealing is indeed nearly completed during the first run so that the curve of $(Q_{\text{ann}}''-Q_{\text{ann}}')_T$ *versus* T enables one to estimate the

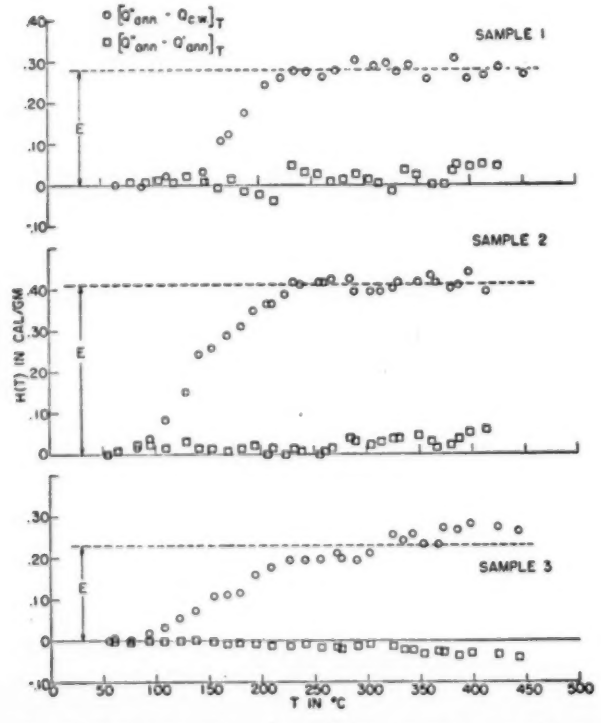


FIG. 5. $H(T)$ *versus* T for three different samples, containing different quantities of stored energy.

TABLE III. The values of E for three different samples as obtained from the data of Fig. 5, together with the corresponding values of W and the resulting latent heat ratios E/W.

Sample	W (cal/g)	E (cal/g)	E/W
3	8.2	0.23	0.028
1	9.4	0.28	0.030
2	11.9	0.41	0.034

[17] Handbook of Chemistry and Physics (Chemical Rubber Publishing Company, Cleveland, Ohio, 1949), thirty-first edition, p. 1816.

error in E to be of the order of 10 percent. It can be seen from these curves that this is so and that in the first run the release of latent energy begins at a temperature of between 100° and 125°C and is practically finished at a temperature of 250°C. As expected, the spread of the points becomes larger as the temperature increases. E, which is obtained as a simple average of the values of $H(T)$, above 250°C, is given in Table 3 together with the value W. The last column shows the values obtained for E/W.

A comparison between Tables 1 and 3 shows that our results, contrary to those of Sato, fail to show any saturation in this range of W so that the ratios of E/W are, as explained before, equal to e/w. On the other hand, the actual values of E/W are much closer to those of Sato than those of Taylor and his co-workers.

It can be seen that the spread in the values of E/W obtained in the present investigation is of the order of magnitude of the variations to be expected on the basis of the errors involved in the annealing experiment. These errors, consequently, represent one of the factors limiting the accuracy of the determination of E/W.

The diameter of the samples and the methods of introducing cold work (twisting) was the same in Quinney and Taylor's investigation as in ours, the only difference being the much faster rate of twisting employed in the former. Whether this fact is responsible for the discrepancy is not known.

The author wishes to express his gratitude to Dr. G. Groetzinger for his continued encouragement and interest in this investigation. In addition, he is indebted to Dr. P. Schwed for suggesting this experiment, and to him and Dr. R. Smoluchowski for valuable discussions; to Dr. J. Burkig and Mr. R. Webeler who participated in designing the apparatus and carrying out the measurements.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 8 AUGUST, 1952

Production of Monochromatic X-Radiation for Microradiography by Excitation of Fluorescent Characteristic Radiation*

T. H. ROGERS
Machlett Laboratories, Incorporated, Springdale, Connecticut
(Received March 3, 1952)

In several types of applications of microradiography, particularly for histochemical and microchemical analysis by x-ray absorption spectroscopy, it is desirable to use monochromatic radiation of a wavelength selected specifically with reference to an absorption edge of the element of interest. An x-ray tube having as its target a material whose K_α characteristic wavelength is suitable is the most usual source of monochromatic x-radiation. In many cases, however, an otherwise desirable target is physically unsuitable for use as a target inside a vacuum tube. Hence, the possibility of producing fluorescent characteristic radiation with intensities sufficient for microradiography becomes of interest. The very high intensity of the primary x-ray beam made available by the newly developed beryllium-window x-ray tubes suggests new possibilities for the excitation of fluorescent radiation.

This paper reports the results of exploratory work attempting to establish the factors governing the intensity and degree of monochromaticity of fluorescent x-radiation. Several materials, ranging from titanium (atomic number 22) to molybdenum (42) were studied comparatively, with exciting radiations of various characteristics; titanium was studied in greater detail. It is concluded that a considerably higher degree of monochromaticity is obtainable by this method than by direct bombardment and that intensities useful for some aspects by microradiography are obtainable.

INTRODUCTION

THE species of microradiography[1] with which this paper will be primarily concerned is frequently designated as "historadiography," which is the term that has been applied to the microradiography of histologic sections. Early development of this art was carried on prior to 1936 by Lamarque,[2] who placed the sections to be radiographed and the film within the evacuated envelope of the x-ray tube itself, so that extremely low energy x-rays might be employed unimpeded by passage through a wall or window material to the tube's exterior. The subsequent development of tubes with thin beryllium windows[3,4] greatly facilitated all such forms of radiography requiring x-rays of relatively long wavelengths, owing to beryllium's relatively high degree of transparency to such radiation.

As originally developed, historadiography employed a continuous spectrum of x-radiation in the long wavelength region between 1.0A and 10A, for which even extremely thin histologic sections have appreciable absorption, and by means of which radiographic contrast between regions of slightly different densities can be made perceptible. This type of spectrum is produced when a conventional x-ray tube is operated with an anode voltage of 12 kv or less.

---

* Presented at Fourth Annual A.I.E.E. Conference on Electronic Instrumentation and Nucleonics in Medicine, at New York (January 7, 1952).

[1] G. L. Clark, *Applied X-Rays* (McGraw-Hill Book Company, Inc., New York, 1940), third edition, pp. 180–184.

[2] P. Lamarque, *L'historadiographie; nouvelle technique d'examen des coupes microscopiques*, Presse med. **44**, 478–481, March 21, 1936; Am. J. Roentgenol. Radium Therapy **33A**, 504–505 (1937).

[3] R. R. Machlett, J. Appl. Phys. **13**, 398–401 (1942).

[4] T. H. Rogers, Ind. Radiography **4**, 35–39 (1945–46).

More recently, a refinement of the historadiographic technique has been developed, whereby the elementary chemical composition and distribution within the cellular structure of the histologic specimen can be determined. The theoretical basis for this method has been well worked out by Engstrom,[5] who has also demonstrated the technical requirements. This "histochemical analysis" requires the use of essentially monochromatic x-ray beams of wavelengths suitably chosen on each side of a characteristic absorption edge of the chemical element of interest. Theoretically, at least, any element except hydrogen can be so analyzed qualitatively and quantitatively in very minute samples, provided monochromatic x-ray beams of suitable wavelengths can be made available.

Workers in this field as well as others requiring monochromatic x-radiation have, in some cases, obtained satisfactory results by using characteristic radiation direct from an x-ray tube having a target of such material as to produce a characteristic line of suitable wavelength, using the filtering techniques commonly employed by x-ray crystallographers to eliminate some of the undesired wavelengths. This approach to the production of "nearly" monochromatic x-rays, while it is the one generally used heretofore, has the restrictive limitation in the choice of wavelength to the characteristic lines of those materials whose physical properties render them suitable for inclusion as the target member in an x-ray tube structure. An alternative method that suggests itself is to excite the desired characteristic spectral lines of the proposed target material by irradiating it with a powerful beam of x-rays of wavelengths shorter than the characteristic wavelength desired. The intensity of this fluorescent radiation is very much lower than that of the incident radiation exciting it, but the development[4,6] in recent years of extremely high intensity x-ray tubes with beryllium windows, which may be used as the primary source for exciting the fluorescent radiation, greatly enhances the possibilities of this approach. Engstrom[7] has reported using this method successfully.

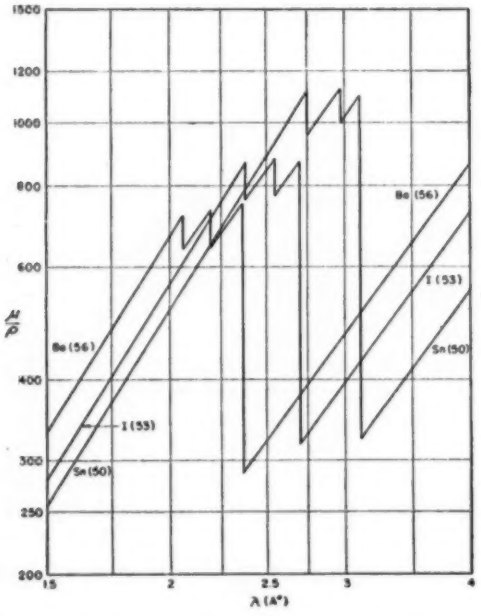


FIG. 1. Variation of mass absorption coefficient (μ/ρ) with wavelength (λ) about the L absorption edges for barium, iodine, and tin.

The investigation reported in this paper represents one phase of a program involving the development of more advanced techniques for historadiography and histochemical analysis, beginning with the work of Marinelli, Ferlazzo, Fitzgerald, and Foote[8] at the Sloan-Kettering Institute. This work involved the analysis of human thyroid tissue for iodine content, by means of historadiography done with a conventional tungsten target tube equipped with a special beryllium window 0.2-mm thick so as to provide efficient transmission of x-rays having wavelengths around 2.7A. Areas of absorption believed to be due in part to iodine content were repeatedly deomonstrated in normal thyroid tissue, and peculiarities in the distribution thereof were correlated with certain pathologic conditions. The identification of iodine on the basis of polychromatic radiation could not be considered positive, however, (for reasons apparent from Fig. 1), and the program was later extended to include the procurement of suitable monochromatic sources to permit the "straddling" technique of Engstrom to be employed for both positive identification and quantitative determinations.

It so happens that in the case of iodine, the deep L_{iii} absorption edge at 2.714A can be properly straddled by the K_α emission lines of titanium at approximately 2.745A and vanadium at approximately 2.50A, respectively. The collaboration of Machlett Laboratories was enlisted in the development of special low voltage, high current x-ray tubes with ultra-thin beryllium windows and titanium and vanadium targets, and in the production of the tubes in sealed-off form.[9] In this connection it was also decided to make some investigation of the possibilities of utilizing secondary fluorescent radiation for this and similar purposes, since if feasible this method would have certain very decided advantages, particularly if it should be desired to analyze for other elements in other parts of the body. Major advantages would be a much wider selection of wavelengths as discussed earlier, elimination of the need for a multiplicity of x-ray tubes, and the removal of limitations imposed at the longer wavelengths by window absorption.

[5] A. Engstrom, Acta Radiologica, Supplementum LXIII (1946).

[6] T. H. Rogers, Radiology 48, 594–603 (1947).

[7] A. Engstrom, Rev. Sci. Instr. 18, No. 9, 681–682 (1947).

[8] Reported at Annual Meeting of Am. Assoc. of Pathologists and Bacteriologists in Boston, April, 1949.

[9] T. H. Rogers and W. B. Lurie, Cathode Press 7, No. 4, 6-9, 30–33 (1950).

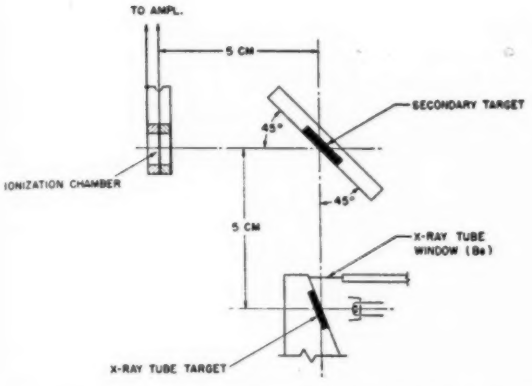


FIG. 2. Diagram of apparatus for studying secondary radiation.

Some of the results of these investigations of secondary radiation phenomena are thought to be of possible interest not only in the field of historadiography but also in other fields of physics concerned with secondary x-radiation. This perhaps overly long introduction has been given so that these results may be presented in their proper perspective against the background of the over-all program of which the investigation was a part. This part of the program was carried out by Machlett Laboratories under the auspices of the Sloan-Kettering Institute and in close cooperation with its Department of Physics. The helpful advice and suggestions of L. Marinelli and G. Ferlazzo are particularly to be acknowledged. Most of the measurements on which the data contained in this report are based were made by W. B. Lurie† at Machlett Laboratories.

EXPERIMENTAL PROCEDURE AND RESULTS

The physical arrangement of the apparatus used in making the measurements of secondary radiation was as indicated in Fig. 2. Except where otherwise stated, the secondary radiating material was in the form of a button 0.75 in. in diameter and 0.1 in. thick, and a beam 3 mm in diameter was permitted to impinge upon the radiator. The radiation was detected, and its relative intensity was measured by means of an ionization chamber and dc amplifier, constituting a direct-reading x-ray intensity meter which has been described elsewhere.[10] The chamber is air-filled; its window and collecting electrode consist of 0.022-mm thick cellulose acetate sheet coated with Aquadag. The indications of this instrument can be expected to be approximately linearly related to roentgens per minute with little correction for wavelength over the range from 1.5 to 2.7A in view of the chamber construction.

In Fig. 3 is shown the intensity of the secondary radiation from copper as a function of the voltage at which the tube constituting the primary x-ray source is operated. The intensity is expressed in roentgens per minute, based on a calibration of the ionization chamber and amplifier detection system at 10 kv by comparison with a Victoreen Nylon chamber r meter and applying the correction indicated by Day[11] at the Bureau of Standards. Also plotted in Fig. 3 is the output of the tube used as the primary source, as measured at a distance of 20 cm from the target. These intensities are shown as roentgens per milliampere of x-ray tube current. At a maximum tube loading of 30 ma at 60 kv, a maximum intensity in the order of 200 r/min is thus obtainable from copper as a secondary radiator.

From this rather crude indication that usable intensities of secondary radiation are obtainable, the next step was to explore how the intensity varies with the atomic number of the secondary target material and also with the target material of the x-ray tube generating the primary radiation. For this purpose, a group of tubes of the type normally used for x-ray diffraction applications, having targets of tungsten, molybdenum, copper, iron, and chromium, respectively, and beryllium windows about $\frac{3}{4}$ mm thick, were employed. Molybdenum, zirconium, copper, iron, chromium, and titanium were used for secondary radiators. Measurements of secondary radiation intensities with the various combinations of primary and secondary targets were made under conditions indicated in Fig. 2, and the results are plotted as five groups of curves, each group consisting of the curves for the various secondary radiators when excited by radiation

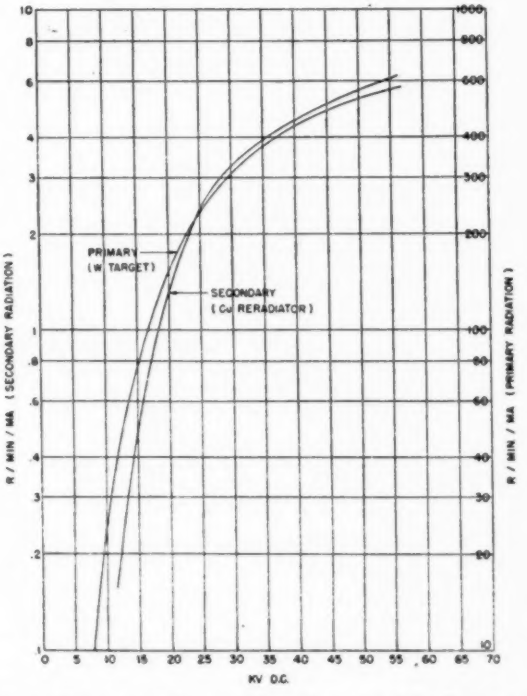


FIG. 3. X-ray intensity as a function of voltage from copper as a secondary target and tungsten as a primary target. (Secondary intensity measured at 5 cm from reradiator placed 5 cm from primary target; primary intensity measured at 20 cm from target.)

† Now with Fleetwood Laboratories, Yonkers, New York.

[10] W. B. Lurie, Electronics, **23**, No. 3, 74 (March, 1950).

[11] F. H. Day, Bur. Standards J. Research, **41**, 295–300 (1948).

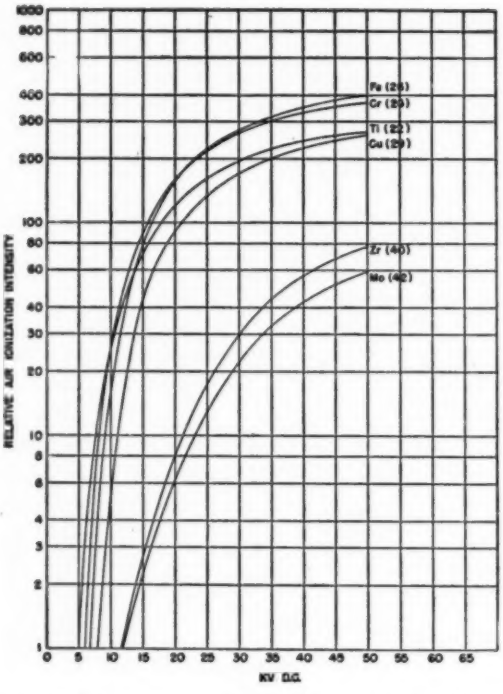


FIG. 4. Relative intensity of secondary radiation from various elements as a function of voltage applied to tungsten target x-ray tube as primary radiation source.

from one of the primary targets. The curves show relative intensity, as measured by the ionization effect, plotted as a function of x-ray tube voltage with constant tube current.

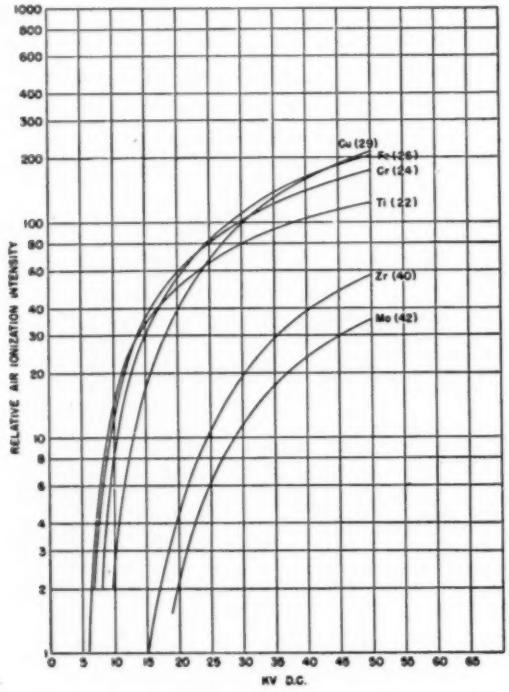


FIG. 5. Relative intensity of secondary radiation from various elements as a function of voltage applied to molybdenum target x-ray tube as primary radiation source.

Figure 4 shows data obtained with the tungsten primary target. Note that the lower atomic number elements begin to radiate appreciably at lower voltage than do the higher atomic number elements. For each of the copper, iron, chromium, and titanium group, appreciable intensity begins to appear at a voltage corresponding approximately to the wavelength of the K-absorption edge of the radiating material. This fact together with the fact that the total secondary radiation is considerably greater from the elements in the lighter group than from the heavier ones, for this limited voltage range at least, suggests the probability that the secondary radiation consists very largely of characteristic fluorescent wavelengths when the incident

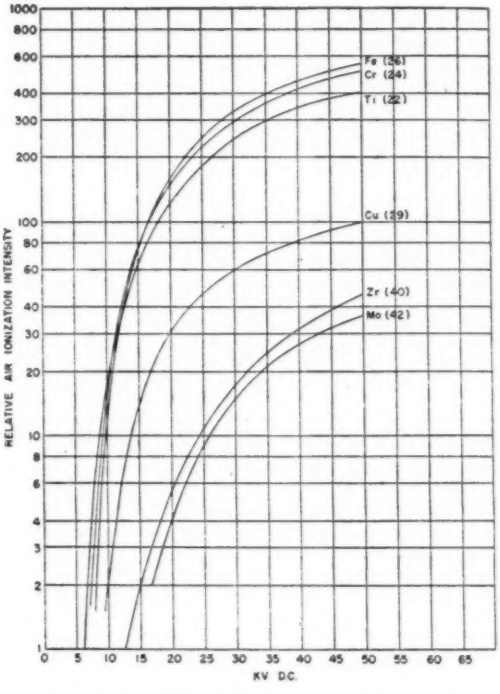


FIG. 6. Relative intensity of secondary radiation from various elements as a function of voltage applied to copper target x-ray tube as primary radiation source.

radiation contains a high proportion of wavelengths not a great deal shorter than the pertinent absorption edge.

Figure 5 shows the data obtained with the molybdenum target tube. The lower intensities, compared with Fig. 4, are probably attributable to the lower efficiency of the molybdenum target as a generator of x-rays. However, the slopes of the curves are greater in the region above 20 kv than for Fig. 4, suggesting that the K-characteristic wavelengths of molybdenum which begin to appear at that voltage enter into the excitation of fluorescence in those elements whose K-absorption edge lies at a wavelength greater than the Mo-K_α. The crossover of the curve for copper, which did not occur in the previous graph, may possibly be explainable by

the occurrence of the Mo-K_α line in the exciting radiation.

Figure 6 shows the results obtained with the copper target tube. Here the effect of the Cu-K_α wavelength concentration is quite striking. The total intensity from the elements whose K-absorption edge is at a longer wavelength than Cu-K_α is greater than ever before; for the others, particularly copper itself, it is lower than before.

Figure 7 shows similar results with reference to the iron target, and Fig. 8 shows similar results with reference to chromium, wherein all intensities are reduced except from titanium which has its absorption edge at a longer wavelength than the Cr-K_α line.

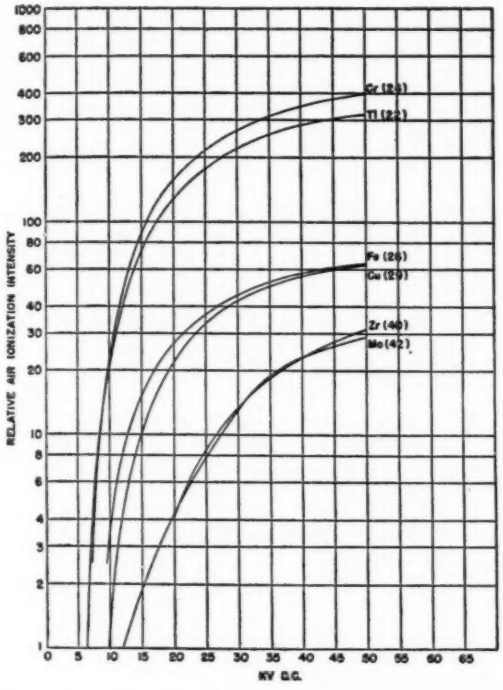


FIG. 7. Relative intensity of secondary radiation from various elements as a function of voltage applied to iron target x-ray tube as primary radiation source.

A close comparison of these five groups of curves reveals some inconsistencies, which may be due to the somewhat crude method of making the measurements, amplifier instabilities, etc. However, it is quite clear that, within the energy range explored, secondary radiation from any given material tends to be produced most efficiently by incident radiation from a target whose K_α line is of a wavelength as near as possible to but shorter than the wavelength of the K-absorption edge of the secondary radiating material. It is also apparent, however, that the total secondary radiation produced by using the tungsten target is, in most cases, almost as intense as when the indicated optimum target is used It would, of course, be much simpler if a single (tungsten target) tube could be used regardless of which secondary radiator would be selected. The

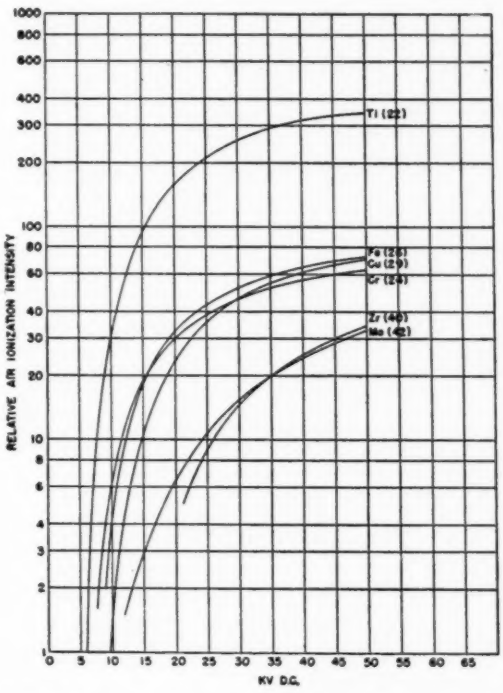


FIG. 8. Relative intensity of secondary radiation from various elements as a function of voltage applied to chromium target x-ray tube as primary radiation source.

matter of the degree of monochromaticity, or "purity" of the secondary radiation, that is, what proportion is characteristic fluorescence and what proportion is Compton scattering, must also be considered.

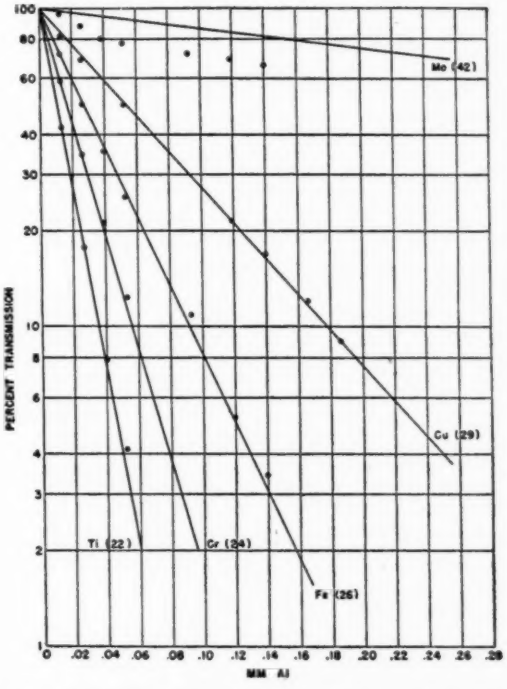


FIG. 9. Aluminum absorption curves for secondary radiation from various elements excited by radiation from tungsten target at 30 kv. (Straight lines are theoretical curves at K_α wavelengths; plotted points are experimentally determined.)

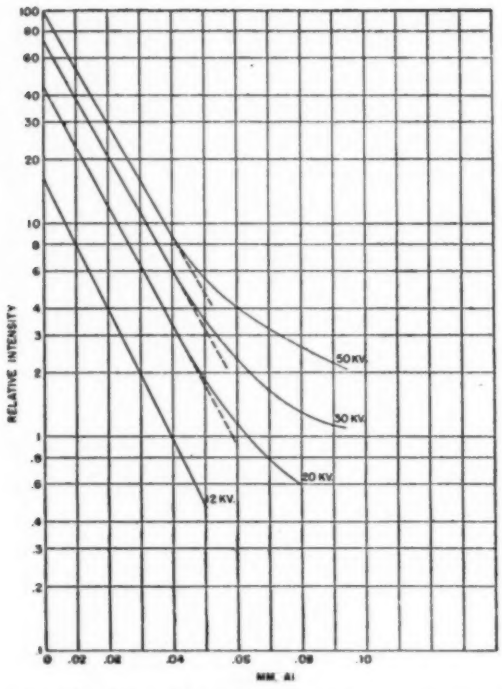


FIG. 10. Aluminum absorption curves for secondary radiation from titanium excited by radiation from tungsten target at various voltages. (Dotted lines theoretical for monochromatic radiation.)

The method of investigating the "purity" was to make aluminum absorption curves and plot them on semilog coordinates. For monochromatic radiation such a plot would make a straight line. Departure from the straight line gives an indication of the "impurity" content of the secondary radiation. Figure 9 shows such plots for titanium, chromium, iron, copper, and molybdenum when irradiated by a tungsten target at 30 kv. The straight lines are drawn with slopes corresponding to the absorption for the wavelengths of the respective K_α's. A high degree of purity is indicated in all cases except for molybdenum, which shows a considerable component softer than the Mo-K_α and then begins to show monochromaticity.

A more intensive study of the purity phase was concentrated on the single element, titanium. Figure 10 gives an indication of how the purity as well as intensity of secondary radiation from titanium varies with the voltage on the tungsten target tube. Using an arbitrary definition of "purity," the percentage of absorption at which the curve departs significantly from the straight line, the purity is above 90 percent at 50 kv and improve as the voltage is decreased. It apparently exceeds 98 percent at 12 kv.

The data plotted in Fig. 11 give an indication of the effect of the thickness of the secondary radiator on both intensity and purity. These particular data were taken with 47 kv on a tungsten target tube, and titanium secondary targets 2.0 mm, 0.017 mm, and 0.0085 mm thick, respectively. Reducing the thickness from 2.0 to 0.017 mm results in a reduction of the total intensity but in an increase in the purity. The further reduction in thickness further reduces the intensity, without appreciable improvement in purity, at that voltage.

By using the 0.017-mm titanium secondary target, a series of measurements were made at a fixed voltage of 25 kv with tungsten, molybdenum, copper, and chromium primary targets to demonstrate the simultaneous effects of primary target material on both intensity and purity of the secondary radiation. These data are plotted in Fig. 12. The chromium target tube produces the greatest intensity of secondary radiation, and it also gives the highest purity. It will be noted that the total intensity with tungsten is about half that obtained with chromium, and the purity comparison is tungsten, about

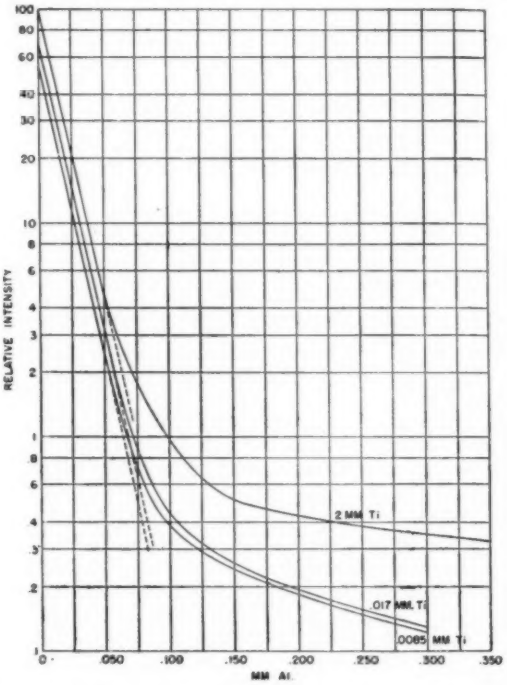


FIG. 11. Aluminum absorption curves for secondary radiation from titanium targets of various thicknesses, excited by radiation from tungsten target at 47 kv. (Dotted lines theoretical for monochromatic radiation.)

98 percent, chromium about 99.7 percent, at this voltage. There would certainly be many cases in which the tungsten tube would be usable as an all-purpose medium as compared with the "optimum," which in this case would be chromium.

Measurements were also made to compare the intensity and purity of the secondary radiation obtained from titanium under approximately optimum conditions with the intensity and purity obtainable directly from the titanium target tubes with ultra-thin beryllium windows that were developed specifically for the historadiographic program. The optimum operating conditions for these tubes were established as 7 kv, 15 ma, employing a 0.017-mm titanium filter to cut out wavelengths shorter than the K_α-characteristic. The titanium secondary target was irradiated by a chromium

target tube operating at 25 kv, 30 ma. Figure 13 shows the results obtained. The intensity of the direct radiation from the titanium target tube was 7.4 times that from the secondary target. The purity from the direct target was about 95 percent; from the secondary, about 99.7 percent. Thus, for the actual operation of historadiography, the time factor would favor the direct by about 10 to 1; precision of results would favor the secondary method, under optimum conditions.

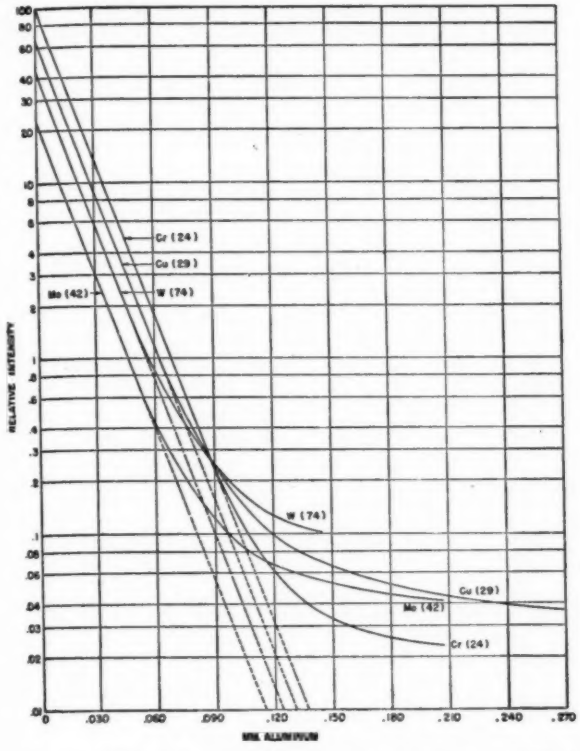


FIG. 12. Aluminum absorption curves for secondary radiation from 0.017 mm thick titanium target excited by radiation from targets of various materials at 25 kv. (Dotted lines theoretical for monochromatic radiation.)

CONCLUSIONS

It has been shown that monochromatic radiation in the wavelength regions pertinent for some aspects of historadiography can be produced by the excitation of characteristic fluorescent radiation in suitably selected elements, with a higher degree of monochromaticity than is usually obtainable directly from x-ray tubes with targets of the same or comparable materials. The intensity so obtainable is sufficient to be usable for these applications.

The investigations made indicate that for a considerable range of energies of incident radiation, the secondary radiation is due to fluorescence to a much greater extent than to Compton scattering. Also, the closer the wavelengths of the incident radiation approach, from the short wavelength side, the *K*-absorption edge of the irradiated material, the more efficient the production of the fluorescent radiation becomes. So far, no attempt has been made to generalize upon these results and relate them to the theory of atomic structure.

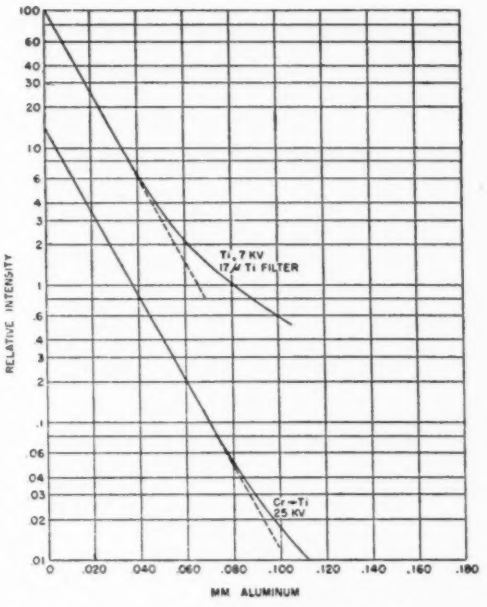


FIG. 13. Aluminum absorption curves for primary radiation from titanium target at 7 kv, 15 ma, filtered by-0.017 mm titanium, and for secondary radiation from 0.017-mm titanium target excited by radiation from chromium target at 25 kv, 30 ma.

Longitudinal and Transverse Ultrasonic Waves in a Synthetic Rubber*

A. W. NOLLE, *Department of Physics, The University of Texas, Austin, Texas*

AND

P. W. SIECK, *Acoustics Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts*

(Received March 24, 1952)

Velocity and attenuation of both transverse and longitudinal ultrasonic waves are measured in a buna-N vulcanizate at 2, 5, and 10 megacycles, over a range of temperatures. The method involves use of solid transmission media to conduct pulsed signals into a thin flat specimen. Values of the bulk modulus and of the complex shear modulus are computed from the measurements. It is found that longitudinal-wave losses in this particular synthetic rubber are substantially accounted for by the relaxation phenomena measured in shear, but this conclusion is subject to considerable quantitative uncertainty arising from experimental errors estimated to be as large as 5 percent for longitudinal-wave acoustic data and 10 percent for transverse-wave acoustic data. It is found that below approximately −25°C the shear-modulus loss tangent becomes relatively independent of temperature and frequency in the present range.

I. INTRODUCTION

THE purpose of this investigation is to extend certain previous studies[1,2] of ultrasonic-wave propagation in rubberlike materials to substantially lower temperatures, and to extend the studies to include transverse waves as well as the usual longitudinal waves.

The large attenuation of ultrasonic waves in rubberlike materials requires that specimens only a few wavelengths in thickness be used for experimental work. A convenient method for dealing with a thin specimen, used in the studies to which we have referred, is to insert the specimen in a liquid bath through which pulsed ultrasonic signals are propagated and to observe the resulting changes in amplitude and transit time of the signal pulses. The velocity and attenuation of ultrasonic waves in the specimen are computed from these observations. The results obtained in this way show that the phase velocity of longitudinal ultrasonic waves in rubberlike materials increases with decreasing temperature, and that the attenuation at constant frequency has a maximum with respect to temperature.

The liquid transmission path used in the above investigations becomes unsatisfactory at low temperatures because of excessive attenuation, and restricts the measurements at ordinary temperatures to longitudinal waves. In the present investigation we attempt to remove these restrictions by introducing a solid transmission path, and to extend the measurements accordingly in the case of a vulcanized specimen of buna-N rubber. In certain ranges of temperature and frequency we obtain data for both longitudinal waves and transverse waves, and with these data we demonstrate (within rather crude limits) that the viscoelastic relaxation phenomena evidenced by ultrasonic longitudinal-wave measurements are attributable to relaxation mechanisms operating in shear rather than in compression.

II. METHOD AND APPARATUS

The electronic apparatus for generating and detecting pulsed signals, and for measuring amplitude and transit time, is functionally identical with that used for pulse-reflection work in previous measurements.[1] In the present application we make provision for connecting the equipment either (1) as a pulse-reflection apparatus for use with a single quartz-crystal transducer, or (2) as a pulse-transmission apparatus for use with separate transmitting and receiving crystals. In the present study the operating frequencies are 2, 5, and 10 megacycles.

The acoustic portion of the apparatus, illustrated in Fig. 1, comprises quartz crystal plates, solid transmission blocks, the rubber sample, and interfacial joint layers. When longitudinal-wave measurements are to be made, X-cut crystals are used in conjunction with aluminum transmission blocks. A thin layer of liquid may be used for each of the interfacial joints. A mixture of silicone stopcock grease with castor oil was frequently used and was found to remain in place over long periods of time. The entire assembly as used for longitudinal-wave work is held together by spring tension. Multiple layers of rubber may be inserted as necessary to obtain a convenient total attenuation at the lower temperatures.

In the case of transverse-wave studies, Y-cut crystals are used in conjunction with Pyrex glass transmission blocks. The glass blocks, while undesirable because of slow attainment of thermal equilibrium, are employed in order to alleviate (1) differential expansion at the crystal joints and (2) an unfavorable acoustical situation,† both of which present difficulty when metal

* A portion of this work was performed at the Acoustics Laboratory, Massachusetts Institute of Technology, with support from Contract NObs-25391, Task No. 1, from the Bureau of Ships; other portions were supported by the Defense Research Laboratory of The University of Texas.

[1] A. W. Nolle and S. C. Mowry, J. Acoust. Soc. Am. **20**, 432–439 (1948).

[2] Ivey, Mrowca, and Guth, J. Appl. Phys. **20**, 486–492 (1949).

† Metal transmission blocks are unsuitable for transverse-wave work by the present method because the approximate two-to-one ratio of longitudinal-wave velocity to transverse-wave velocity, characteristic of a number of metals, allows a longitudinal wave

blocks are used in transverse-wave work. The crystal-to-block joints are thin layers of ceresin wax. The block-to-rubber joints consist of a resin layer, in contact with the block, followed by a layer of modified rubber in contact with the sample. The entire joint, which is formed from commercial cement preparations, is cured under a pressure of about 50 lb/in.^{2} at a temperature somewhat above 100°C.

The time required for an echo to return from the sample interface to the crystal is measured for each transmission block in turn. Then the A to B transit time is measured. The difference between the time required for a signal to travel from crystal A to crystal B, and the average of the echo times, represents the time required for the signal to traverse the sample. Similarly, the difference between the relative level in decibels of the transmitted signal and the average of the echo levels, when corrected by well-known relations for partial reflection at the interfaces, represents attenuation due to the sample. The velocity in cm/sec and the attenuation in db/cm are easily computed when the sample thickness is known. The interfacial reflection correction depends, to a good approximation, upon the densities and the wave velocities in the two media, provided that the interfacial joint is so thin that it may be considered to have infinite impedance. The velocity in the transmission blocks is computed from values of the time difference between successive echoes and of the lengths of the blocks.

The assembly of Fig. 1 is immersed in a well-stirred alcohol bath in an open-top Dewar flask. Suitable housings protect the crystals and their associated electrical networks. Temperature is controlled by adjusting the flow rate of liquid nitrogen through a heat exchanger contained in the bath, and by adjusting the current through an electrical heating element. We estimate that the temperature of the specimen is known to ±0.5°C.

A timing error of a fraction of a microsecond may be introduced by failure to maintain exactly constant electrical impedances when the apparatus is changed over from the pulse-reflection connection to the pulse-transmission connection. In the case of longitudinal-wave measurements, we correct for this error, when it is found to exist, on the basis of velocity measurements made by the liquid-path method.[1] In the case of transverse-wave measurements, we guard against serious error of this nature by performing a corroborative measurement of the velocity with an interference method. This method, which can be used only at the lower temperatures where attenuation is not excessive,

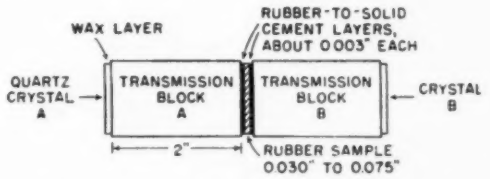


FIG. 1. Acoustic portion of apparatus arranged for introducing ultrasonic signals into a rubber specimen by transmission through solids.

involves adjusting the velocity in the specimen (by regulation of the temperature) to find minima and maxima in the transmission loss. These represent the well-known conditions for resonance and antiresonance for plane waves in a path terminated by high impedance plane reflecting surfaces. Attenuation values reported in final data are taken, however, only under conditions where large total attenuation in the specimen makes interference effects negligible.‡

The manner in which the sample must be incorporated in the system for transverse-wave measurements may be expected to introduce appreciable systematic error. The interfacial joints associated with the sample, whose acoustic properties are not accurately known, have a total thickness from 10 to 20 percent of the sample thickness. Furthermore, the sample is processed under heat and pressure as already described. In addition we expect random errors up to 10 percent in the transverse-wave velocity measurements, because the large values of the transverse-wave attenuation require the use of thin samples. We have no absolute means for evaluating the total error in transverse-wave velocity results. The disagreement found in results for different samples and different frequencies suggests that the total error is in the range 10 to 20 percent.

In the temperature ranges in which data are available for comparison, the longitudinal-wave velocity and attenuation results agree within 5 percent with values later obtained by an improved set of liquid-bath measurements not yet published, except that over part of the temperature range the timing error already described gave rise to a more severe disagreement in the 5-mc velocity data. In this temperature range the 5-mc velocity data were replaced by liquid-bath results, as will appear in the plots of data. The timing error becomes relatively unimportant for longitudinal-wave measurements at the lower temperatures, where we found it possible to add layers of rubber up to a total thickness of 0.8 cm.

Our rubber specimen is a vulcanized buna-N compound which we designate as B-100. The composition in parts by weight is Hycar OR-15, 100; ZnO, 5; S, 1.5; benzothiazyl disulfide, 1.3. The specimen is cured under pressure for 60 minutes in a mold at 143°C. This compound is considered to be of interest because the

which is twice reflected within one block to reach the second crystal at nearly the same time as the directly transmitted shear wave. The Y-cut quartz crystal in practice produces a small longitudinal-wave signal along with the transverse signal. At the higher temperatures, where the transverse-wave attenuation is very large, this undesired signal may be the stronger of the two after both signals have traversed the sample. When glass transmission blocks are used the two signals are well separated in time.

‡ An apparent secondary maximum of attenuation with respect to temperature, which appears in one of the graphs of reference 1, is in fact attributable to interference.

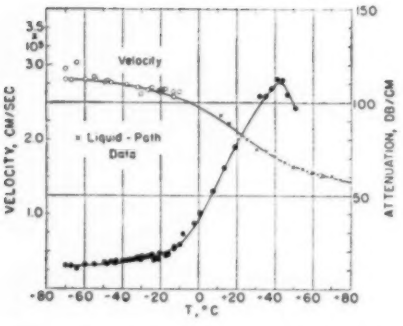


FIG. 2. Velocity and attenuation of 5-megacycle longitudinal ultrasonic waves in buna-N vulcanizate B-100.

temperature of its attenuation maximum is relatively high, so that we can extend measurements to temperatures far below that of the maximum. The compound was used also in some previous studies at audio and ultrasonic frequencies.[1,3]

III. RESULTS

In Figs. 2 and 3 we show graphically the longitudinal-wave results for 5 and 10 mc. The 5-mc velocity data have been augmented with later results by the liquid-path method, as stated above. We note especially that at low temperatures the attenuation is relatively insensitive to temperature and is roughly proportional to frequency. Longitudinal-wave measurements were not made at 2 mc by the present method.

Figure 4 contains the results of the transverse-wave measurements. While the velocity data for the various frequencies define distinguishable curves in some regions, these curves do not show a consistent trend with increasing frequency. We conclude that the apparent differences associated with frequency are due primarily to the systematic errors already described, and that the present experimental procedure is inadequate to resolve the frequency dependence of the velocity. Therefore we have constructed a single velocity-temperature curve which represents the concensus of the data for the three frequencies, and to which we shall refer in analyzing the data.

The assumption that the frequency dependence of shear-wave velocity is smaller than experimental uncertainty in the present ranges of frequency and temperature, which we have made in drawing a single velocity curve, is justifiable by extrapolated results of measurements at lower frequencies on a buna-N compound. (These extrapolated results are found in Fig. 1 of reference 3.) It is also noteworthy that in Figs. 2 and 3 above the velocity of longitudinal waves at 5 and 10 mc is the same within experimental error, except possibly for temperatures above the attenuation peak.

The data of Figs. 2–4 are not corrected for thermal expansion of the sample, but are computed for the room-temperature thickness. The thermal expansion correction is less than 2 percent for temperatures above −25°C.

IV. INTERPRETATION OF RESULTS

Our object in securing data for both longitudinal and transverse waves on the same specimen is to ascertain whether the losses observed in longitudinal-wave measurements can be accounted for in terms of relaxation processes occurring in pure shear. This question cannot be answered from longitudinal-wave data only, since the velocity and attenuation of longitudinal ultrasonic waves depend upon the viscoelastic behavior of the specimen in both compression and shear. Our data provide information for both wave types over a limited range of temperature, but we are able to give only a crude answer to the question stated above because of the uncertainties in the transverse-wave results.

We assume that, for a single frequency, the mechanical behavior of the specimen in shear can be expressed by a complex rigidity modulus

$$\mu^* \equiv \mu_1 - i\mu_2 = \mu_1 - i\eta\omega \tag{1}$$

and by a complex bulk modulus

$$B^* \equiv B_1 - iB_2 = B_1 - i\zeta\omega, \tag{2}$$

where η is the dynamic shear viscosity, ζ is the dynamic compressional viscosity, and ω is the frequency in radians/sec. We shall let c_t, c_l, denote the phase velocities and α_t, α_l, denote the amplitude attenuation constants for transverse and longitudinal waves, respectively. In order to connect the complex modulus quantities with the observed velocity and attenuation, we note that if $M^* \equiv M_1 - iM_2$ is the effective stiffness for a certain wave motion whose velocity is given by the usual relation

$$c^* \equiv c_1 - ic_2 = (M^*/\rho)^{\frac{1}{2}}, \tag{3}$$

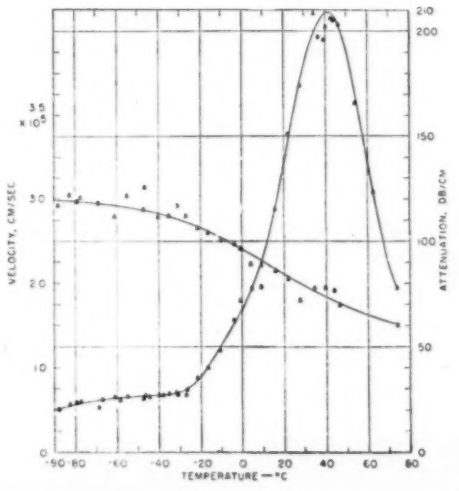


FIG. 3. Velocity and attenuation of 10-megacycle longitudinal ultrasonic waves in buna-N vulcanizate B-100.

[3] A. W. Nolle, J. Polymer Sci. 5, 1–54 (1950).

where ρ is the density, then direct solution of (3) gives

$$M_1=\rho(c_1^2-c_2^2), \tag{4a}$$

$$M_2=2\rho c_1 c_2. \tag{4b}$$

Also, examination of the usual complex exponential expression for a progressive wave shows that c_1, c_2 are related to the observed phase velocity c and the amplitude attenuation constant α by the equations

$$c_1=c/(1+r^2), \tag{5a}$$

$$c_2=rc/(1+r^2), \tag{5b}$$

where $r\equiv\alpha c/\omega$ and ω is the radian frequency. Combination of Eqs. (4a) through (5b) gives the working relations

$$M_1=\rho c^2(1-r^2)/(1+r^2)^2, \tag{6a}$$

$$M_2/M_1=2r/(1-r^2). \tag{6b}$$

Since the complex stiffness for transverse waves is μ^*, the quantities M_1, M_2 represent μ_1, μ_2 when α_t, c_t are used in (6a), (6b). Since the effective complex stiffness for longitudinal (bulk) waves is $B^*+(4/3)\mu^*$, M_1 and M_2 represent respectively $(B_1+4/3\mu_1)$, $(B_2+4/3\mu_2)$, when α_l, c_l are used in these relations. Evidently B_1, B_2 can be evaluated only when velocity and attenuation data are available for both transverse and longitudinal waves.

We are interested primarily in finding out whether the viscoelastic losses in shear are sufficiently large to account for the longitudinal-wave attenuation. Therefore we shall calculate, in addition to certain other dynamic quantities, the quantity

$$F\equiv(4/3)\mu_2/[B_2+(4/3)\mu_2], \tag{7}$$

which is the fraction due to pure shear of the viscoelastic losses computed from longitudinal-wave data. A value of unity for F would indicate that longitudinal-wave attenuation is due entirely to shear losses.

Values for the velocity and attenuation of longitudinal and transverse waves at various frequencies and temperatures, where both wave types were studied, are given in Table I, together with certain values of the viscoelastic quantities computed from the acoustic data by relations (6a), (6b). Attenuation results are given in decibels/cm and are therefore 8.69 times as great as the corresponding α-values. The 2-mc longitudinal-wave data are taken from later unpublished results obtained by an improved technique with the liquid-path method; other acoustic data are taken from the smoothed curves of the charts already presented. All acoustic data have been corrected for thermal expansion of the sample. Density values were obtained by combining a directly measured value of 1.05 g/cm³ at 23°C with the volume thermal expansion coefficient of 0.0012/C°, which was computed from variation of the thickness of one sample over a wide temperature range. The 5-mc values of longitudinal-wave velocity were used in the calculations.

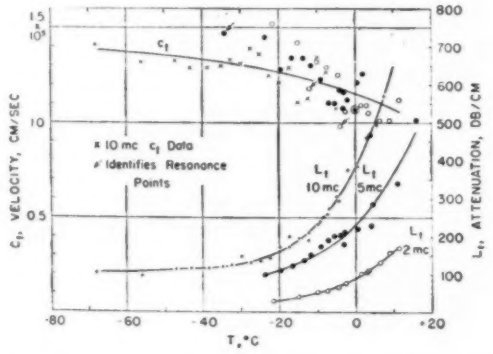


FIG. 4. Velocity and attenuation of transverse ultrasonic waves in buna-N vulcanizate B-100 at 2, 5, and 10 megacycles. A single velocity curve is used to represent the data from the three frequencies, because the expected dependence of velocity on frequency is less than experimental error in the temperature range covered.

TABLE I. Ultrasonic data for buna-N compound B-100.

		Centigrade temperature							
		−65	−50	−30	−20	−10	0	10	15
Density, g/cm³		1.16	1.15	1.12	1.11	1.09	1.08	1.07	1.06
Wave velocity, units 10^5 cm/sec	Transverse	1.33	1.32	1.27	1.25	1.20	1.14	1.06	1.03[a]
	Longitudinal (5 mc)	2.57	2.54	2.51	2.46	2.41	2.34	2.23	2.16
Transverse-wave attenuation, db/cm	2 mc				31.5	52.7	88.8	162	
	5 mc				113	155	233	367	459
	10 mc	108	115	130	162	237	385	646	
Longitudinal-wave attenuation, db/cm	2 mc						15.1	21.5	25.7
	5 mc	12.6	13.8	17.0	19.1	24.9	37.8	57.5	68.2
	10 mc	25.9	28.0	28.6	36.1	52.7	69.1	92.4	113
μ_1, real part of shear mod., units 10^{10} d/cm²		2.06	2.00	1.81	1.72	1.56	1.36	1.14	1.02
B_1, real part of bulk mod., units 10^{10} d/cm²		4.84	4.79	4.61	4.46	4.23	4.03	3.70	3.46
μ_2/μ_1, Shear loss factor	2 mc				0.072	0.12	0.19	0.35	
	5 mc				0.10	0.13	0.19	0.29	0.35
	10 mc	0.053	0.056	0.061	0.074	0.11	0.16	0.25	
F, fraction of longitudinal-wave losses arising in shear	2 mc						0.88	1.13	
	5 mc				1.00	1.02	0.93	0.88	0.89
	10 mc	0.78	0.77	0.79	0.77	0.74	0.84	0.95	

[a] Extrapolated value.

Discussion of Compressional Behavior

The "F" values of Table I vary from 0.74 to 1.13; since these values range around unity, the longitudinal-wave losses are accounted for at least roughly by the dynamic shear viscosity. We now inquire how much the "F" values might vary from unity as a result of experimental error alone, the supposition being made momentarily that the longitudinal-wave losses do in fact arise in shear and that therefore correct measurements should always give a value of unity for "F". From Eqs. (6a), (6b) we find the approximate relation

$$F\simeq(4\alpha_t c_l^3)/(3\alpha_l c_l^3), \tag{8}$$

TABLE II. Ultrasonic data from reference 6 for "Polymer D," mole wt. 5560. Frequency 4.5 mc.

T, °C	Density,[a] g/cm³	Long.-wave attenuation, nepers per wavelength[b]	Long.-wave velocity, cm/sec[c]	Shear-wave surge resistance,[d] g/sec-cm²	Shear-wave surge reactance,[d] g/sec-cm²
10	0.899	0.637	1.98×10^5	6.7×10^4	2.1×10^4
40	0.886	0.425	1.53×10^5	3.03×10^4	1.8×10^4

[a] Reference 6a, Fig. 3. [b] Reference 6a, Fig. 11. [c] Reference 6a, Fig. 12. [d] Reference 6c, Fig. 14.

which is valid for $r\ll1$ and which is accordingly reasonably accurate for present data. From (8) we find that the lowest observed value for F, which is 0.74, could arise from a fortuituous combination of velocity and attenuation errors of less than four percent, the attenuation and velocity values being too small for transverse waves and too large for longitudinal waves. Alternatively, use of the standard relations for combination of probable errors (not necessarily applicable to the present experiments, where much of the error may be systematic) shows that the observed "F" values lie within the range of probable error when the probable error in the transverse-wave quantities is less than 10 percent and the probable error in the longitudinal-wave quantities is less than 5 percent. We conclude that the "F" values of Table I, in spite of the relatively wide range of variation, are equivalent to unity within present experimental error. As a corollary, it is evident that a high degree of precision in the acoustical measurements is necessary to lead to reasonable close values of the complex bulk modulus.

In each case where data are available for computation of an "F" value, the same data allow us to compute a value of the bulk-modulus loss factor B_2/B_1. These results, not shown in the table, range from small negative values (attributable to uncertainty in the data) to positive values less than 0.02. Thus, although because of experimental uncertainty we cannot assert that bulk losses are entirely negligible as a source of longitudinal-wave attenuation, we do find definitely in the range explored that the loss factor ("loss tangent") for the bulk modulus is much less than that for the shear modulus (μ_2/μ_1), which ranges from 0.053 to 0.35.

Similar small values for B_2/B_1 have been reported by Meyer and Taum[4] and by Kuhl and Meyer[5] on the basis of results obtained at audio and lower ultrasonic frequencies. Values of the order of 0.03 are reported for certain rubbers. The method depends on the effect of a rubber slab upon resonances or standing waves in a liquid system. These authors also report values of 0.08 to 0.15 for specimens measured by ultrasonic absorption at 10 mc, but since there is no indication of an independent study of transverse waves, it must be concluded that these latter values were based on longitudinal-wave measurements alone and represent $[B_2+(4/3)\mu_2]/[B_1+(4/3)\mu_1]$ rather than B_2/B_1. In this case the latter data agree with room-temperature values for the buna-N specimen of the present study.

Useful comparisons may be made with the results of Mason and associates,[6] who have performed ultrasonic investigations of polyisobutylene liquids by the transmission of longitudinal waves and also by the reflection of transverse waves (the reflection occurring within a quartz block from a flat face coated with the liquid). In order to make comparisons of results, one must take into account some differences in terminology. In reference 6, "compressional viscosity" refers to the viscosity associated with that Lamé elastic constant which is generally denoted by λ in treatises on elasticity, whereas in the present paper we understand by "compression" the all-sided contraction which is associated with the bulk modulus, and the loss term B_2 is associated with such volmetric deformation.§ Also "compressional stiffness" in reference 6c corresponds to the present $B_1+4\mu_1/3$. It appears more fruitful in discussion of results to use bulk modulus and shear modulus rather than the Lamé constants, since the former two modulus quantities, taken individually, have direct experimental significance and often arise naturally in theory, whereas the λ Lamé constant is a derived quantity lacking such direct significance.

In reference 6 it is found that the viscosity associated with the λ Lamé constant is negligible in the case of a polyisobutylene liquid of molecular weight of about 1000; for other polyisobutylene liquids, with molecular weights in the range 3000 to 6000, this viscosity in the range of several megacycles is at least as large as the high frequency shear viscosity. The viscosities just mentioned are frequency-independent in the special sense that they correspond to values of elements in equivalent circuits which are shown, in reference 6, to represent quite accurately the dynamic properties of the various polyisobutylene liquids. For present purposes a more direct comparison is afforded by making calculations of B^* and μ^* directly from the

TABLE III. Values of dynamic quantities from data of reference shown in Table II.

T, °C	μ_1 Real part dyn. shear mod., d/cm²	μ_2 Im. part dyn. shear mod., d/cm²	μ_2/μ_1	B_1 Real part dyn. bulk mod., d/cm²	B_2 Im. part dyn. bulk mod. d/cm²	B_2/B_1	F
10	0.50×10^{10}	0.32×10^{10}	0.64	2.7×10^{10}	0.27×10^{10}	0.10	0.61
40	0.068×10^{10}	0.12×10^{10}	1.80	2.0×10^{10}	0.12×10^{10}	0.060	0.59

[4] E. Meyer and K. Taum, Akust. Zeits. **7**, 45–50 (1942).

[5] W. Kuhl and E. Meyer, "Dynamical properties of rubber and rubberlike materials in a large frequency range," in *Noise and Sound Transmission*, 1948 (Physical Society, London, 1949), pp. 181–188.

[6] (a) Mason, Baker, McSkimin, and Heiss, Phys. Rev. **73**, 1074–1091 (1948); (b) **74**, 1873–1874 (1949); (c) **75**, 939–946 (1949).

§ The conversion relation $B_1=\lambda+\frac{2}{3}\mu_1$ may be employed; a relation of the same form applies to the imaginary parts of the quantities. The symbol λ is employed also to denote certain viscosity quantities in reference 6. Note that a negative value for the viscosity counterpart of λ is required for $B_2=0$.

acoustic data in these references. Such data are shown in Table II for one polymer at a single frequency and two temperatures. The quantities calculated from these data are shown in Table III. These calculations are made by use of Eqs. (6a), (6b) with longitudinal-wave propagation data, and by use with the shear-wave surge impedance data of the following relation:

$$\mu^* \equiv \mu_1 - i\mu_2 = (Z_{\text{surge}}^2)/\rho. \tag{9}$$

The "F" values of Table III are less than any found from the data found in the present paper and must be considered as significantly less than unity, for the longitudinal-wave velocity measurements by Mason and associates are estimated as correct within one percent[6a] (more precise than corresponding results in the present paper) and the other quantities of Table I appear to be measured with precision at least somewhat better than that obtained for the solid samples of the present work. Moreover, the values of B_2/B_1 in Table III for the polisobutylene liquid are much greater than those found from present data for vulcanized buna-N rubber.

Evidently then, our conclusion (subject to considerable uncertainty), that the dominant longitudinal-wave losses in buna-N rubber are not those associated with all-sided compression, is not necessarily a general statement for all high polymers. A possible explanation reconciling the two findings (but one which could be tested only by much experimentation) is that the temperature dependence of bulk relaxation mechanisms is different from that for shear relaxation mechanisms, and that the relationship of the two temperature dependences varies with different polymers. According to this point of view, the measurements of reference 6 happened to be taken in a temperature region where both relaxation mechanisms are important in polyisobutylene, whereas the present work happened to fall in a temperature region far from the bulk loss peak for buna-N.

In general some skepticism is warranted with regard to compressional (bulk) loss data. Inhomogeneities in the specimen, in particular small bubbles, can give rise to losses which are macroscopically indistinguishable from compressional losses, but which are in reality due to shearing deformation in the regions of the inhomogeneities. Also in experiments where a small specimen is supposed to be subjected to all-sided compression in an acoustically excited liquid,[4,5] the observed losses may include a contribution due to shear unless the specimen is small compared to a wavelength in all dimensions and is subjected to acoustic waves of accurately uniform phase and amplitude. There is often practical difficulty in realizing the latter requirement.

The real part of the bulk modulus B_1 increases regularly with decreasing temperature (Table I). This is a normal feature of organic liquids, other than water, and is not necessarily indicative of a relaxation phenomenon. The fractional increase in B_1 per degree of temperature decrease, for the buna-N specimen, is in fact somewhat less than corresponding values which we have found in other measurements for methyl-ethyl ketone, which is a solvent for the buna-N preparation. In Table III we observe that B_1 for the polyisobutylene liquid, studied by the Mason group, increases steeply with decreasing temperature, although the absolute values are considerably smaller than those of Table I for the buna-N specimen.

Discussion of Shear Behavior

We observe values for μ_1, the real part of the dynamic shear modulus, of roughly 1 to 2 times 10^{10} d/cm^2. It is interesting that these values, obtained with a buna-N vulcanizate having average molecular weight of at least several hundred thousand, are at comparable temperatures only about twice as large as those obtained by Mason and co-workers for polyisobutylene of viscosity-average molecular weight 5560 (see Table III and reference 6a). That group showed,[6c] however, that further increase of the molecular weight of polyisobutylene to 100,000 increases the shear modulus by only about 13 percent (at 10 megacycles, 25°C). The general principle suggested here, that high frequency dynamic behavior for a given polymer is relatively independent of molecular weight, had been asserted previously on theoretical grounds.[7] There is of course some numerical difference in the high frequency shear modulus for the buna-N preparation and for polyisobutylene of high molecular weight; this must be explained in terms of intrinsic differences between the polymers as regards (a) relaxation times and (b) values of the ultimate "crystalline" modulus which corresponds to infinite frequency.

Finally we call attention to the behavior of the longitudinal wave attenuation in the buna-N specimen, as shown in Figs. 2 and 3. We have concluded that this attenuation represents largely shear losses, at least in a range of temperatures below 15°C. According to the above graphs, the attenuation is approximately proportional to frequency; since for temperatures well below the attenuation peak the real parts of the modulus quantities are relatively independent of frequency, it follows from Eq. (6b) that for such temperatures the quantity μ_2/μ_1 is substantially independent of the frequency. A corresponding result was obtained for polyisobutylene liquids by the Mason group,[6] who pointed out that at higher frequencies (above a few megacycles at room temperature) the shear behavior of their specimens can be represented by a "hysteresis element"—a viscoelastic element having a phase angle independent of frequency.

Also in the data of Figs. 2 and 3 we observe that below some −25°C the attenuation becomes relatively insensitive to temperature, so that μ_2/μ_1 becomes

[7] See, for example, Turner Alfrey, *Mechanical Behavior of High Polymers* (Interscience Publishers, Inc., New York, 1948), pp. 145–148.

relatively insensitive to temperature as well as frequency. This result appears in Table I. The temperature −25°C is approximately the upper limit of the "glass-like" range for buna-N rubber as determined, for example, by brittleness tests.

ACKNOWLEDGMENT

We acknowledge the cooperation of Mr. J. F. Mifsud of The University of Texas Defense Research Laboratory, who provided certain of the longitudinal-wave data.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 8 AUGUST, 1952

Space Charge Reduction of Low Frequency Fluctuations in Thermionic Emitters

T. B. Tomlinson
Department of Electronics, University College, Southampton, England
(Received March 17, 1952)

The methods of measuring the space charge reduction factor for shot noise are briefly reviewed, and attention is drawn to the difficulties of making a similar measurement for flicker noise. An indirect qualitative test is described in which a triode is used both in normal manner and as a diode under retarding field conditions. Noise measurements at a high and a low frequency show that the effect of space charge on flicker noise exceeds that for shot noise. The diode results are used to estimate the value of the flicker noise reduction factor which is then seen to have a value less than the corresponding shot noise reduction factor at all current levels.

INTRODUCTION

IT has been known since the first report by Johnson[1] in 1925 that thermionic emitters generate an anomalous fluctuation noise in excess of shot noise at low frequencies. Johnson's results were subsequently examined theoretically by Schottky[2] who introduced the term "flicker effect." The state of knowledge today is still little in advance of that at the time of Schottky's paper. Several theoretical approaches[3–6] have been made to this and to the related fluctuation effect in semiconductors, but up to the present, insufficient experimental evidence has been obtained for a full verification of these theories. It is probable that further understanding of the effect may follow from a close comparison of the properties of flicker noise with those of shot noise, about which a great deal more is known. Some of the essential differences are already established; for instance, the mean square noise current $\langle \delta i_f \rangle_{Av}^2$ due to flicker effect in the narrow frequency interval δf, centered on frequency f, appears to obey a law of the type

$$\langle \delta i_f \rangle_{Av}^2 = K \frac{I^m}{f^n} \cdot \delta f,$$

where $m \doteq 2$, $n \doteq 1$. K is not a true constant; it varies from valve to valve, is dependent on the working conditions, etc. The mean square shot noise current $\langle \delta i_s \rangle_{Av}^2$ in a temperature limited diode for the same frequency interval is given by the well-known Schottky formula

$$\langle \delta i_s \rangle_{Av}^2 = 2eI \cdot \delta f,$$

where e is the electronic charge. If the fluctuations are investigated by means of frequency selective apparatus having response $\phi(f)$, the measured mean square noise current will have a value

$$\langle \delta i_s \rangle_{Av}^2 = 2eI \int_0^\infty \phi(f) \cdot df = 2eI \cdot \Delta f,$$

where Δf may be referred to as the integrated band width of the apparatus.

The shot noise fluctuations are of purely statistical nature and do not depend on the properties of the emitter itself; hence, the formula is correct at all frequencies below those at which transit time effects enter. When the valve is operated under normal working conditions a negative space charge is established in the interelectrode region, giving effect to a potential minimum. (See Fig. 1.) The presence of this minimum plays an important role in reducing the fluctuations of space current below the theoretical value. As a result a reduced magnitude of the mean square noise current is measured, and this may be conveniently expressed in the form

$$\langle \delta i_s \rangle_{Av}^2 = \Gamma^2 \cdot 2eI \cdot \Delta f,$$

where $\Gamma^2 < 1$ and is termed the "space charge reduction factor." The theoretical evaluation of Γ^2 has interested many authors,[7–10] and considerable experimental in-

[1] J. B. Johnson, Phys. Rev. **26**, 71 (1925).
[2] W. Schottky, Phys. Rev. **28**, 74 (1926).
[3] G. G. Macfarlane, Proc. Phys. Soc. (London) **59**, 366 (1947) and **63**, 807 (1950).
[4] A. Van der Ziel, Physica **16**, 359 (1950).
[5] J. M. Richardson, Bell System Tech. J. **29**, 117 (1950).
[6] F. K. du Pré, Phys. Rev. **78**, 615 (1950).
[7] D. O. North, R.C.A. Rev. **4**, 441 and **5**, 106 (1940).
[8] A. J. Rack, Bell System Tech. J. **17**, 592 (1938).
[9] W. Schottky and E. Spenke, Wiss. Veroff. Siemens-Werken **16**, 1 (1937).
[10] F. C. Williams, J. Inst. Elec. Engrs. (London) **88**, Part III, 219 (1941).

vestigation of the space charge reduction of shot noise has been carried out. On the other hand, little attention has been given to the measurement of a corresponding reduction factor Γ_f^2 for flicker noise currents. Schottky has expressed the opinion,[11] based on theoretical grounds, that the reduction effect of space charge on flicker noise currents will exceed that for shot noise. The following is an account of experimental measurements designed to investigate the differences between the effects of space charge on flicker noise and on shot noise fluctuations.

OUTLINE OF METHOD

There exists no theoretical formula for the mean square value of the unreduced flicker noise current; therefore, it is necessary to measure both this and the space charge limited noise. The most direct method of measuring shot noise under the two different conditions is now described. The apparatus required consists of a diode valve with an anode load, a high gain amplifier tuned to a suitable frequency, and a thermocouple with galvanometer for measuring the mean square value of the output noise voltage from the amplifier. The frequency chosen must be high enough for flicker effect to be negligible but sufficiently low for transit time effects to be unimportant. Measurements are made under two conditions:

(a) The filament is at normal temperature, and the anode voltage is adjusted to give a certain specified anode current, so that the valve works under space charge limited conditions. In order to calculate the noise current of the diode a correction must be made for noise voltages introduced by the amplifier and for thermal noise in the anode load. The shunting effect of the dynamic resistance r_a of the diode must be taken into account, as will be shown later.

(b) The anode supply voltage is raised and the filament current reduced until the original value of anode current is obtained, which is now independent of the anode voltage. Under these conditions (saturated or temperature limited) all the emitted electrons proceed to the anode, and the potential minimum, with which the space charge reduction of noise is associated, has disappeared from the interelectrode space (Fig. 1). The new mean square output voltage corresponds to unreduced shot noise, i.e., $\Gamma=1$. The new root mean square noise current is computed, with correction for extra noise sources. The ratio of the diode noise currents under the two conditions gives the space charge reduction factor.

Complete temperature saturation is not achieved in valves using oxide-coated cathodes; therefore, the above experiment gives precise results only with a tungsten emitter. In the case of flicker noise, it has been known since Johnson's original work that oxide-coated and tungsten cathodes differ in their behavior. Furthermore, experiments show that whether space charge is present or not the magnitude of the flicker noise current depends on the temperature of the cathode. For these reasons the above method cannot be used when examining flicker noise at low frequencies.

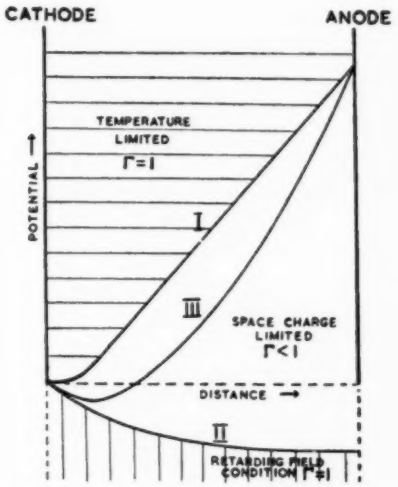


FIG. 1. Potential distribution in a parallel-plane diode illustrating the three modes of operation. Curve I shows the point of lowest potential at the cathode; reduction of cathode temperature causes entry into the temperature limited region. Curve II shows the lowest potential at the anode; the diode is at critical current and a more negative anode voltage causes entry into the retarding field conditions. Curve III shows a typical potential distribution under normal working conditions. The potential minimum appears in the interelectrode space and thus plays an important role in the reduction of noise fluctuations.

To make such a direct comparative measurement it would appear that the same cathode surface at a fixed temperature must supply the same space current both in the absence of space charge and with space charge present. This is, of course, impossible. However, the difficulty may be avoided by the following indirect method. A triode is selected as the test valve and is first used with anode and grid tied together. The amplifier is tuned to a low frequency, and the noise current is measured with the test valve working as a diode at some specified current under "retarding field conditions." A more detailed description of this condition is given below.

When the anode voltage of a diode is progressively reduced from zero to negative values, the decrease of anode current is accompanied by a modification of the potential minimum between cathode and anode. Below a certain critical anode current the minimum has disappeared and the point of lowest potential is at the anode (Fig. 1); hence, no space charge reduction of noise is to be expected. This condition has been extensively studied in papers[12,13] by Fürth and MacDonald in which it is shown that the shot noise current is accurately given by the Schottky equation, i.e.,

[11] E. B. Moullin, *Spontaneous Fluctuations of Voltage* (Clarendon Press, Oxford, England, 1938), Ch. 6, p. 165.

[12] D. K. C. MacDonald and R. Fürth, Proc. Phys. Soc. (London) **59**, 375 (1947).

[13] R. Fürth and D. K. C. MacDonald, Proc. Phys. Soc. (London) **B63**, 300 (1950).

$\Gamma=1$ for all currents below the critical current. Consequently, the low frequency noise current measured under these conditions may be taken to be the unreduced noise current. It is convenient to express its mean square value $\langle \delta i_L \rangle_M^2$ in terms of the unreduced shot noise for the same current, *viz.*,

$$\langle \delta i_L \rangle_M^2 = F_D^2 \cdot 2eI\Delta f,$$

where F_D is a factor applicable to the valve connected as a diode. It is to be remembered that $\langle \delta i_L \rangle_M^2$ contains both flicker noise and shot noise components so that F_D^2 is always greater than unity.

After the measurement of noise for the diode connection, the test valve is reconnected as a triode. A high positive anode supply voltage is used, and the negative grid bias voltage is adjusted to give the same anode current as before. At the low current concerned the large negative grid bias will produce an electron-optical effect dividing the electron flow into small streams passing between the centers of grid apertures, and emission from the cathode regions under the shadow of grid wires will cease. This implies that the electron density in the region of emitting areas is relatively high so that some space charge effect may be expected. A second noise measurement is made under these conditions and a new factor F_T applicable to the triode connection is calculated. The ratio F_D/F_T is a measure of the space charge reduction effect on the low frequency fluctuations, introduced by changing to the triode connection. The two noise measurements are repeated at a much higher frequency in order to obtain a corresponding ratio Γ_D/Γ_T for shot noise fluctuations at the same specified current. Comparison of the two ratios F_D/F_T and Γ_D/Γ_T indicates qualitatively the relative effects of space charge on low and high frequency fluctuations. For all the valves tested it was found that the ratio F_D/F_T reached a maximum value at a current in the region of 100 μa. Although space charge had set in for the diode above its critical current, the effect of space charge in the triode was still the greater of the two. Consequently, it was possible to extend the range of measurement to currents considerably in excess of the critical current. An extrapolation method involving the results for F_D below the critical current enables an estimate of Γ_f to be made.

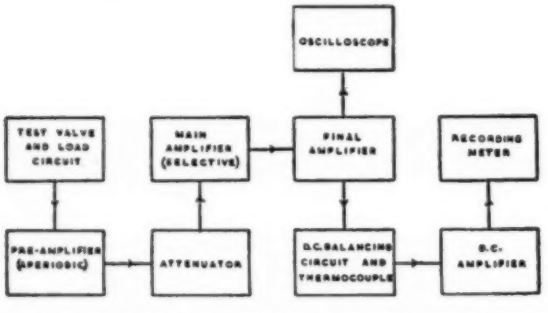


FIG. 2. System used in measuring the fluctuations of current in a valve at low frequencies.

Three different methods of measuring the rms value of the noise voltage at the anode of the test valve are known. In the first of these, the noise voltage is compared with a small sinusoidal voltage introduced into the anode circuit from a signal generator. This method requires accurately attenuated signal voltages of the order of a fraction of a microvolt, and it also requires knowledge of the integrated bandwidth of the apparatus. In the second method the signal generator is replaced by a tungsten diode operated in the temperature limited condition so that its rms noise current is calculated directly from the anode current. Unfortunately, this elegant method cannot be employed at low frequencies because of the flicker noise originating in the comparator diode itself. In the third method the noise voltage is determined by comparison with the calculated thermal noise voltage across a resistor connected in the anode circuit of the test valve. In this and the second method, the band width need not be measured.

The third procedure was the one here adopted; no extra resistor was introduced, the thermal noise in the anode load circuit itself serving as a source of comparison. It is extremely important that no "current" noise should be generated in this circuit and a specially tested wirewound resistor was employed.

MEASUREMENTS AT THE LOW FREQUENCY

Figure 2 is a block diagram of the system; a more detailed diagram of the test valve circuit and the first stage of the two stage preamplifier is given in Fig. 3. The usual precautions were taken against microphony and regenerative feedback. High voltage supplies for the first stages were obtained from dry batteries, and adequate decoupling guarded against the introduction of extraneous noise from this source. The filament supply voltage for the test valve and for the two preamplifier valves was obtained from a large capacity 6-volt storage cell which, after an initial period of discharge, could be relied upon to give a steady filament voltage over a long time interval. When the test valve was triode-connected, the grid impedance had to be kept low to avoid the introduction of new noise voltages, e.g., thermal noise, noise due to gas current, etc. The two preamplifier valves were operated at a low current[11] ($Ia=100$ μa) in order to keep their noise contribution small compared with that in the test circuit. These valves were specially chosen from a large number for their low flicker noise. All resistors used in this section were either of the wirewound or high stability type in order to eliminate current noise. For the same reason all capacitors were of the paper dielectric type and selected for extra low leakage current. A small percentage of commercial wirewound resistors develop a high noise when passing dc current; therefore, all the resistors were specially tested for the effect before use.

[11] D. A. H. Brown, T.R.E. Memo. No. 329 (unclassified, 1950).

The main amplifier consisted of two units employing selective twin-T networks similar to a type described elsewhere.[15] It was chosen for its stability of gain and frequency response. The band width of the two units was 1.7 cps for a mid-band frequency of 20 cps. This frequency was sufficiently low for the magnitude of the flicker noise to be measurable in the presence of the shot noise even at the low operating currents under retarding field conditions. The ratio of band width to mid-frequency is small enough for the flicker noise spectrum to be considered uniform over the pass-band.

The final amplifier was simply a single stage of low gain employing negative feedback and capable of delivering a large peak output voltage. This output voltage was fed into a cathode follower circuit, the thermocouple filament being connected between the cathode output terminal and the cathode of a similar dc balancing valve. This arrangement was adopted to overcome the difficulty of ac connection at the low frequency, the resistance of the filament being only 1500 ohms. The output of the thermocouple was fed in push-pull to a balanced dc amplifier with cathode follower output circuits and with the pen-recording milliammeter connected between the two cathodes. Full scale deflection was obtained for a dc input of 2 mv. As the thermocouple delivered a maximum output of 7 mv, this ensured that it was operating as a square law device. The combination of final amplifier, thermocouple, and recorder was calibrated as a whole for ac signals. Long-time constant circuits were incorporated in the dc amplifier to reduce the short-term fluctuations of the current in the recording milliammeter. The mean square value of this current is derived from the recording by visually averaging the deflections over a long-time interval.

For a given accuracy, the minimum recording time is inversely proportional to the over-all band width of the apparatus.[16] The average recording time was about 10 minutes; results of particular importance were recorded over intervals as long as 30 to 40 minutes. Most readings were taken on three separate occasions to ensure that the results obtained were repeatable. On occasions anomalous readings were obtained, but these occurred mostly at the beginning of a series of measurements before the equipment had completely settled down. It was found advisable not to record deflections before a stabilizing period of at least one hour had elapsed. This waiting period also allowed the drift in the balancing circuit and dc amplifier to decrease to negligible proportions. The monitor oscilloscope was employed to make certain that no extraneous disturbances were being included in the measurement. As a further precaution the work was carried out in a doubly screened room with filters in all supply leads.

[15] J. M. Sturtevant, Rev. Sci. Instr. **18**, 124 (1947).
[16] R. R. Bennett and A. S. Fulton, J. Appl. Phys. **22**, 1187 (1951).

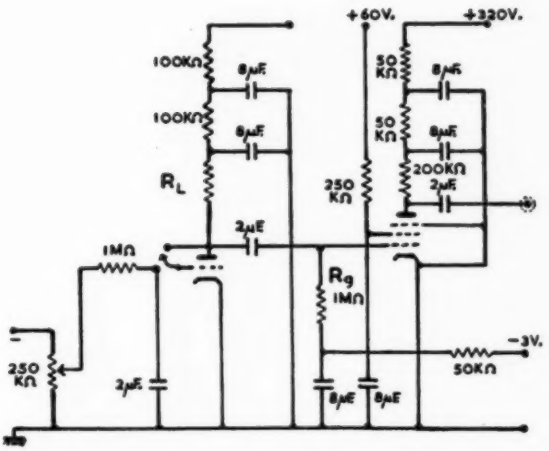


FIG. 3. Circuit diagram of the test valve associated load circuit, and the first stage of the preamplifier as used for the low frequency measurement.

With the valve acting as a diode, three noise measurements were required:

(a) The input to the preamplifier was shorted, the main amplifier gain was set to a suitable point, and the mean square output noise voltage V_{sc}^2 was recorded. This reading is due to the inherent noise in the preamplifier and in subsequent stages.

(b) The short circuit was removed, and the new output voltage V_v^2 was recorded.

Now

$$V_v^2 = V_{sc}^2 + R'^2\{\langle \delta I \rangle_{Av}^2 + \langle \delta I_R \rangle_{Av}^2\},$$

where R' represents the resistance of the load circuit R_L, the grid resistor R_g, and the test valve r_a, all in parallel. $\langle \delta I \rangle_{Av}^2$ is the mean square noise current from the test valve; $\langle \delta I_R \rangle_{Av}^2$ is the thermal noise current in the parallel combination of R_L and R_g.

Put

$$\langle \delta I \rangle_{Av}^2 = F_D^2 \cdot 2eI \cdot \Delta f \quad \text{and} \quad \langle \delta I_R \rangle_{Av}^2 = (4kT_R \Delta f)/R,$$

where $R = (R_L R_g)/(R_L + R_g)$, then

$$V_v^2 - V_{sc}^2 = R'^2\left(F_D^2 \cdot 2eI\Delta f - \frac{4kT_R \cdot \Delta f}{R}\right), \qquad (1)$$

where k is Boltzmann's constant and T_R is the absolute temperature of the load circuit. The power dissipation in this circuit was extremely small; therefore, T_R was taken to be room temperature.

(c) The filament supply to the test valve was next interrupted, and the output voltage V_{oc}^2 recorded. In this case $\langle \delta I \rangle_{Av}^2$ and r_a are not effective, so that

$$V_{oc}^2 - V_{sc}^2 = 4kRT_R \cdot \Delta f. \qquad (2)$$

Eliminating Δf from Eqs. (1) and (2) we have

$$\frac{V_v^2 - V_{sc}^2}{V_{oc}^2 - V_{sc}^2} = \frac{R'^2(F_D^2 \cdot 2eI + 4kT_R/R)}{4kRT_R},$$

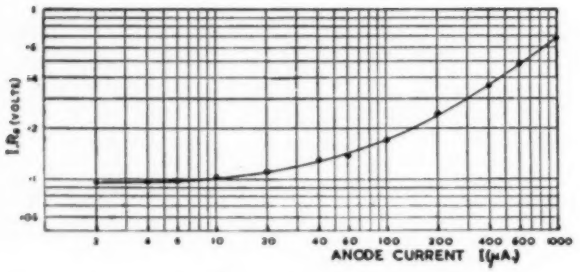


FIG. 4. Examination of test valve characteristic when operated as a diode in the retarding field region. The product Ir_a is constant for currents below the critical current.

giving

$$F_D^2=\left\{\frac{V_v^2-V_{sc}^2}{V_{oc}^2-V_{sc}^2}\cdot\frac{R^2}{R'^2}-1\right\}\cdot\frac{2kT_R}{eIR}.$$

The test valve was reconnected as a triode, the anode supply voltage was fixed at 250 volts, and the negative grid bias voltage was adjusted to restore the anode current to the original value. After a settling down period the three noise voltage measurements were repeated, and the factor F_T was evaluated. Results for F_D and F_T were taken over a current range 0–200 μa.

MEASUREMENTS AT THE HIGH FREQUENCY

In order to keep circuit modifications to a minimum, a frequency of 4 kcps was chosen. If this frequency is too high, the stray capacitance leads to an incorrect assessment of the thermal noise originating in the anode current of the test valve. To ensure that this was not the case at 4 kcps three values of R_L were used (250 kohm, 100 kohm, 50 kohm), and the results were compared. A greater difference between V_v and V_{sc} and, consequently, improved accuracy was obtained by operating the preamplifier valves at a higher current (Ia=0.8 ma). The coupling capacitors were reduced in value to ensure a good cutoff on the low frequency side of the pass-band which was now approximately 320 cps. In this measurement it was not necessary to use a recording meter. The output from the thermocouple was therefore connected directly to a millivoltmeter with thermal scale calibration giving an immediate indication of the mean square output voltage. Results for Γ_D and Γ_T were obtained over the same range as for the low frequency measurements.

THE MEASUREMENT OF THE DYNAMIC RESISTANCE r_a OF THE TEST VALVE

For convenience this measurement was conducted at 4 kcps, use being made of the amplifier and output arrangements described above. The amplifier gain was reduced until the noise output voltage was completely negligible. A small sinusoidal voltage from an audio oscillator was then introduced into the anode circuit of the test valve, and the signal output voltage $\tilde{V}_v$ was observed with the test valve in circuit and operated under the appropriate working conditions. The filament supply was then interrupted, and the new output voltage $\tilde{V}_{oc}$ was observed.

Then

$$\tilde{V}_v=\tilde{V}_{oc}\cdot r_a/(R+r_a),$$

whence

$$r_a=R\cdot\tilde{V}_v/(\tilde{V}_{oc}-\tilde{V}_v).$$

For accuracy, input signal voltages of the order of a few millivolts only were used. In addition, the magnitude of R was made comparable with that of r_a, a low R_L being used for the diode and a high R_L for the triode connection. The calibrated attenuator was used to prevent overloading and to make the output voltages approximately equal under the two conditions.

The product Ir_a was graphed against current I, and corrected values of r_a were read off from the smooth curve. The curves for the diode connection are particularly interesting; an example is shown in Fig. 4. The product Ir_a is constant for all currents up to the critical current above which it increases in value. For all test specimens the critical current was found to be approximately 5 μa.

DISCUSSION OF RESULTS

Results were taken for a number of test valves, all of the same basic type but of different construction and supplied by different manufacturers. Figure 5 shows a specimen set of curves depicting the variation of F_D, F_T, Γ_D, and Γ_T with mean anode current I. It is seen that Γ_D is approximately unity up to the critical current, confirming the observations of Fürth and MacDonald and at the same time proving the method to be reasonably accurate. F_D is greater than unity at low currents because of the flicker noise component; with increasing current its value at first rises, the rise of flicker noise with current exceeding the drop in the total noise due to the space charge effect. With further increase of current, the space charge effect more than compensates for the rise of flicker noise due to increased current so that the value of F_D reaches a maximum and then begins to fall (see also Fig. 7).

The curves of Fig. 5 were used to plot the variation of the ratios F_D/F_T and Γ_D/Γ_T, with current shown in Fig. 6. The ratio F_D/F_T is seen to exceed Γ_D/Γ_T at

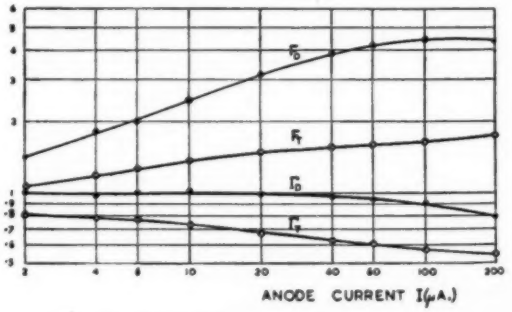


FIG. 5. Variation of the factors F_D, F_T, Γ_D, Γ_T with anode current.

all current levels. Although the magnitude of the flicker noise fluctuations was found to vary between specimen valves, the ratio F_D/F_T was, without exception, the larger of the two. Consequently, the following conclusions may be drawn:

(a) Flicker noise fluctuations of a given current are reduced by the presence of space charge.

(b) This reduction exceeds the corresponding reduction of shot noise for the same conditions of operation of the valve.

ESTIMATION OF ${}_f\Gamma_D$

The observations of F_D were extended to a current of 1 ma for three test valves. Figure 7 shows the results for one particular specimen. As F_D is a combination of both flicker noise and shot noise components, the part ${}_fF_D$ originating solely as flicker effect may be evaluated.

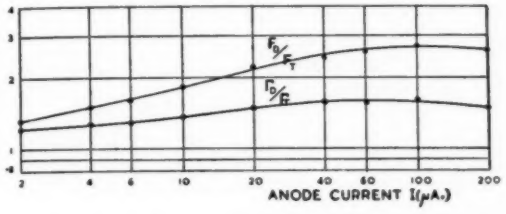


FIG. 6. Variation of the ratios F_D/F_T and Γ_D/Γ_T with anode current showing the effect of space charge on the low frequency noise fluctuations to exceed that for shot noise.

Assuming no correlation between the two noise components:

$$ {}_fF_D{}^2 = F_D{}^2 - \Gamma_D{}^2. $$

The graph of ${}_fF_D$ is seen to be linear for currents below the critical current; by extrapolation of this portion of the curve the magnitude of the unreduced flicker noise at higher currents may be estimated. From this estimate and from the measured value of ${}_fF_D$ a figure for the flicker noise space charge reduction factor ${}_f\Gamma_D$ is derived. Values for ${}_f\Gamma_D$ obtained in this manner are given in Fig. 7, from which it is clear that ${}_f\Gamma_D < \Gamma_D$ for all current magnitudes up to 1 ma. From the trend of the curves this would also appear to be the case at higher currents. For the three test specimens the extrapolated linear curve was found to lie close to the curve ${}_fF_D \propto (I)^{\frac{1}{2}}$, i.e., $\langle \delta i_f \rangle_{\Delta f}{}^2 \propto I^2$, and this has been followed in Fig. 7. It seems certain, however, that this law is not particularly accurate over a wide current range; specimen valves displaying high flicker noise at normal operating currents were found, in some cases, to have a noise level below the average at 2 μa. As Γ_D was measured at 4 kcps only, a second set of measurements was taken at 28 kcps. It is evident from Fig. 7 that flicker effect was not negligible at 4 kcps for the higher currents, but the difference between ${}_f\Gamma_D$ and Γ_D is sufficiently great to confirm Schottky's opinion.

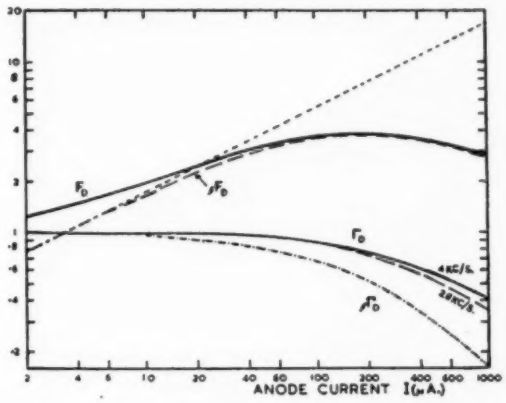


FIG. 7. Variation of F_D, Γ_D, and the flicker noise factor ${}_fF_D$ with anode current. The unreduced value of ${}_fF_D$ is estimated by extrapolating the linear part of the curve at lower current levels. The derived flicker noise reduction factor ${}_f\Gamma_D$ is less than the shot noise reduction factor Γ_D at all current levels.

A previous report by Nicoll[17] describes a similar procedure, but in his case only three experimental points for a single specimen diode formed the basis of the extrapolation and, hence, no claim was made for the accuracy of the derived results. Nevertheless, they are comparable with those of Fig. 7. The three points were those for the total noise; therefore, the flicker noise component in the particular specimen must have been exceptionally high.

Since the reduction of flicker noise by space charge has been established one may conclude that a form of low frequency partition noise exists in multicollector valves. As in the case of shot noise there will be a partial cancellation of the space charge effect due to the distribution of current between the anode and other collector electrodes. Preliminary measurements at a low frequency show quite clearly that the anode current of pentodes includes extra fluctuations due to partition; a full account of this investigation and of the results will be published later.

[17] G. R. Nicoll, T.R.E. Memo. No. 147 (unclassified, 1949).

Electron Microscopy of Synthetic Latices

SAMUEL H. MARON, CARL MOORE, AND ALLEN S. POWELL
Physical Chemistry Laboratory, Department of Chemistry and Chemical Engineering, Case Institute of Technology, Cleveland, Ohio

(Received March 19, 1952)

A study has been made by electron microscopy of the size and size distribution of 22 synthetic latices. Included are GR-S latices Types II, III, and V, polystyrene latices, and several latices of special character.

Details are given on sample preparation, measurement, application of corrections, calculation of number, volume to surface, and molecular weight average diameters, as well as the number and weight average distributions.

The weight average particle size range covered was 730 to 3920A. The smallest sizes were shown by Type II latices, followed in order by Type III, polystyrene, Type V, and some of the special latices. The latices were all nonhomogeneous in size and showed size distributions both on a number and weight basis which in some instances contained single and in some cases double maxima.

INTRODUCTION

SYNTHETIC latices are colloidal dispersions of polymer particles in water. The dispersions are generally stabilized by agents of various kinds, such as fatty or rosin acid soaps or synthetic emulsifiers. The particles themselves may be quite rigid, as in the case of polystyrene, or plastic, as in butadiene-styrene copolymers of relatively high butadiene content. Their size may range all the way from below 100A to several microns in diameter, and depends upon the polymerization recipe, the emulsifier, and the preparation conditions employed.

Study of latex systems by electron microscopy presents a number of difficulties. First, plastic suspended particles exhibit a tendency to flatten during drying when mounted on support films by ordinary techniques. Second, the polymer particles frequently have a scattering power for electrons not much greater than that of the organic substrate upon which they are supported, giving rise thus to poor contrast. Third, samples of commercial synthetic latices consist as a rule of particles quite heterogeneous in size, and hence reliable sampling is important. Finally, small particles may approach the limit of resolution of the microscope, and hence they may not show up distinctly on a micrograph.

To circumvent some of these difficulties, an extended study was undertaken to find a satisfactory and reliable means of preparing micrographs of synthetic latices and to employ such micrographs for determination of particle size and size distribution. After prolonged trial a method was finally evolved which proved satisfactory and which allowed determination of particle size and size distribution of such latices with considerable certainty. This paper describes the micrograph preparation techniques employed and gives also some of the results obtained for 22 synthetic latices of various types.

EXPERIMENTAL DETAILS

Methods of mounting latex samples for electron microscope study have been described by Kelsey and Hanson,[1] Hanson and Daniel,[2] Brown,[3] and Turkevich and Hillier.[4] The technique of Kelsey and Hanson involves imbedding latex particles in polyvinyl alcohol films. Experiments with this method gave micrographs of only fair resolution, and indicated considerable flattening of the particles. The flattening could not be reduced appreciably by previous vulcanization of the rubber as recommended by Hanson and Daniel. However, the flattening could be markedly reduced and the resolution definitely improved by first stiffening the particles by bromination and then mounting them on collodion, as suggested by Brown.

By use of an objective aperture in the microscope a considerable increase in contrast between background and very small latex particles was achieved. Finally, shadowing of the specimen mounts with gold was tried. This technique of Williams and Wyckoff[5] resulted in further improvement of resolution and overcame difficulties with blurred edges on small particles and widely varying density in particles having the same apparent diameter. In addition, it provided a means of checking flattening through measurement of the shadow lengths.

1. Sample Mounting

In preparing specimens for particle size study, latex samples were first diluted by adding one drop or so of latex to 30 or 40 ml of water containing a dispersing agent such as Daxad 11. The actual extent of dilution varied, and was dependent upon the type of latex, its solids content, and the particle size. If, for instance, a latex had a large particle size, a larger quantity of latex was necessary. The sample was then brominated by placing a drop of the diluted latex on top of a small glass pedestal and blowing bromine vapors over it.[6] Here again experience determined the amount of

[1] R. H. Kelsey and E. E. Hanson, J. Appl. Phys. **17**, 675 (1946); Rubber Chem. Tech. **20**, 602 (1947).
[2] E. E. Hanson and J. H. Daniel (private communication).
[3] W. E. Brown, J. Appl. Phys. **18**, 273 (1947).
[4] J. Turkevich and J. Hillier, Anal. Chem. **21**, 475 (1949).
[5] R. C. Williams and R. W. G. Wyckoff, J. Appl. Phys. **15**, 712 (1944); **17**, 23 (1946).
[6] E. H. Rowe (private communication).

bromine required by each sample or type of latex. A portion of the drop of brominated latex was next transferred to a collodion film and allowed to dry in a desiccator. Following the drying, excess soap and dispersing agent were washed off the collodion with hot, freshly distilled water, and the film redried.

The support films were made by casting collodion (Mallinckrodt "Parlodion") from a 2 percent amyl acetate solution on the surface of clean water. The water and film-forming solution were kept free of contamination by frequently renewing the water and cleaning the dropping pipette used for the collodion solution. Portions of the collodion film transferred from the water surface to woven stainless steel screens of 200 mesh served as the mounts for the latex specimens.

Use of the above technique without proper precautions can lead to agglomeration and contamination of the samples, and to excessive flattening of plastic particles. To prevent agglomeration latices must be highly diluted in presence of fairly high concentrations of dispersing agents. Again, to prevent both agglomeration and contamination the water used must be freshly distilled. It is also essential that the mounts be dried in a dust-free atmosphere. The excessive flattening of plastic particles, when observed, is generally the result of incomplete bromination. To make sure that bromination is complete it is advisable to contact the latex with bromine for a reasonable period of time, or to expose it to sunlight or ultraviolet light in order to accelerate the bromination reaction. Finally, too high a concentration of bromine is also to be avoided, for the bromine may attack the collodion film and reduce its strength.

2. Gold Shadowing

Gold shadowing of the specimen mounts was carried out in a RCA vacuum chamber. The gold was deposited on the collodion-wire screen mounts at an angle so as to give a four-to-one ratio of shadow length to transverse particle diameter. This procedure improved considerably the sharpness of particle outline, allowed ready differentiation of latex particles from background, and made measurement of diameter considerably easier and more reliable.

In an effort to eliminate prominence of background from shadow cast collodion, another mounting technique was tried. It involved deposition of a latex specimen on a glass slide, gold shadowing on the glass, and subsequent removal of the entire mount by stripping off with collodion or Formvar films.[7,8] The method proved unsatisfactory, however, since it was very difficult, during the operations of drying, washing, redrying, and shadowing, to retain a surface on the glass from which the thin collodion or Formvar films could readily be stripped.

[7] R. C. Williams and R. C. Backus, J. Appl. Phys. **20**, 98 (1949).
[8] R. C. Williams and R. W. G. Wyckoff, Science **101**, 594 (1942).

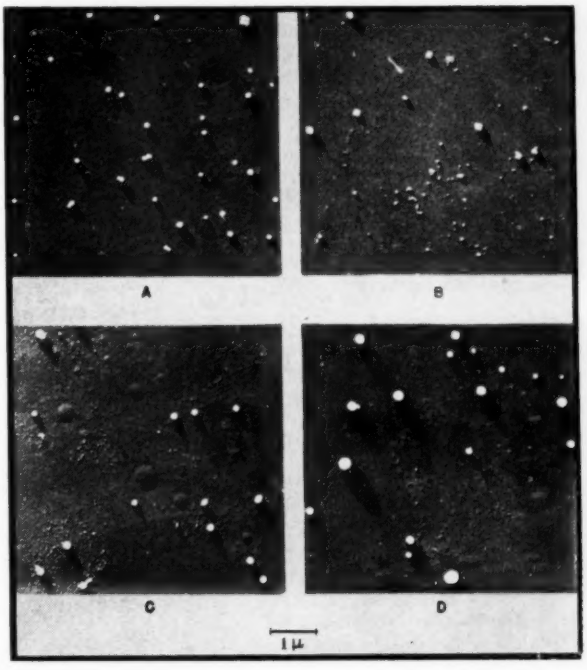


FIG. 1. Electron micrographs of synthetic latices. The samples shown are: (A) Type II GR-S, (B) Type III GR-S, (C) polystyrene latex, and (D) Type V GR-S. (6,000×).

3. Preparation of Micrographs

Micrographs of latex specimens were made with a Type EMU-2 RCA electron microscope using Eastman Kodak contrast emulsion lantern slide plates. To minimize sintering of the gold, the instrument was operated with a low intensity unbiased electron gun. The magnifications used were generally from 3600 to 7500 diameters for the original micrographs and 14,400 to 30,000 diameters for the final enlarged pictures. The magnifications were varied according to the particle size of the latex.

To check the magnification of the electron microscope a simple calibration was carried out. A wire mesh support similar to one used for specimen mounts was punctured and some of the wires broken. By examination of the damaged screen through a calibrated light microscope the distance between the ends of broken wires was carefully measured. The screen was then transferred to the electron microscope, observed at the lowest magnification, and the known distance between the ends of the wires used as a means of checking the magnification. Higher magnifications were checked against this initial calibration.

Typical micrographs obtained in the present study are shown in Fig. 1. For particle size and shadow length measurement positive prints, in which the particles appeared black and the shadows white, were used rather than the negative prints.

4. Measurement of Particle Size and Corrections

For determination of particle size several micrographs of a latex were first prepared. The enlarged prints were

TABLE I. Corrected particle diameter averages for pure soap Type II latices.

Myristate latex No. 235					Laurate latex No. 236				
Print No.	No. of particles	Average diameters-A D_N	D_s	D_w	Print No.	No. of particles	Average diameters-A D_N	D_s	D_w
A-1-1	350	496	685	868	A-1-1	434	407	594	757
A-1-2	290	509	715	824	A-1-2	356	393	565	754
A-1-3 / A-2-2	390	572	778	879	A-1-3	429	413	550	675
A-2-1	285	578	737	830	A-2-2	217	390	565	757
A-2-3	365	488	646	750					
Total or weighted average	1680	531	720	835	Total	1436	404	567	733

then ruled off into convenient segments and the particles in each segment measured and counted. To improve the precision and ease of measurement, a Bausch and Lomb plate viewing magnifier having a 0.1-mm enclosed scale was used, and the photographic enlargements were mounted on a plate viewer fitted with an opal glass illuminated from the rear.

Two corrections, both operating to decrease the observed particle sizes, are required to obtain true diameters. The first is that due to flattening. Latex particles, even when brominated, are flattened somewhat during drying, and so yield diameters that are too large. Although gold was deposited on specimen mounts of both brominated and unbrominated latex so as to give a theoretical 4:1 ratio of shadow length to particle diameter, measurements of shadow lengths on any one micrograph showed that the ratios of the transverse diameters of the particles to the shadow lengths were not constant. In fact, the ratio was not constant even for unflattened particles, except over a limited area, due to slight sagging of the collodion between wires of the supporting screen.

The procedure adopted for determining the flattening in a given micrograph was to select from different areas of the picture particles with the largest shadow length to transverse diameter ratios l/D, and to use the average value of these, T, for estimation of the heights of other particles in the micrograph. From a consideration of the geometry involved, the height D_h of any particle having a shadow length l is readily shown to be

$$D_h = l/T. \tag{1}$$

Assuming now no imbedding of the particles in the collodion film, and that in flattening a spherical particle becomes an ellipsoid, equation of the expression for the volume of the original sphere to that of the ellipsoid gives for the original diameter D_0 of the flattened particle,

$$D_0 = (D_h D^2)^{\frac{1}{3}}. \tag{2}$$

If the flattening correction f be defined as

$$f = D_0/D, \tag{3}$$

then combination of Eqs. (1), (2), and (3) yields

$$f = (D_h/D)^{\frac{1}{3}} = (l/DT)^{\frac{1}{3}}. \tag{4}$$

In practice the flattening factor f was generally determined for about 25 of the largest particles in a field, and the average of these values taken as the correction factor for the entire micrograph. With brominated rubber latex particles the correction generally amounted to less than 10 percent and was never found to exceed 15 percent. With polystyrene latices it was always less than 5 percent.

There are some indications that the latex particles imbed slightly in the collodion substrate. However, it is impossible to differentiate imbedding from flattening, nor is it certain whether this imbedding is uniform for all particles. We believe, however, that the above manner of handling the flattening correction accounts very nearly for both imbedding and flattening effects, and hence no additional correction for imbedding is necessary.

The second correction arises from the fact that on bromination polymer particles undergo a 10 percent increase in diameter,[3,9] thus making the observed size too large by this amount. This correction of 10 percent was applied to all the flattening corrected results given in this paper except those for the polystyrene latices. The polystyrene particles were found to possess sufficient difference in density from the substrate and to be rigid enough to be mounted on collodion without appreciable flattening, and hence no bromination of the polystyrene latices was necessary. Subsequent work on polystyrene and other latices involving the determination of particle size by another method, to be described in a separate publication, confirmed the validity of this correction.

REPRODUCIBILITY OF RESULTS

Several types of particle size averages were computed from the measurements. These included (a) the number average diameter D_N, namely,

$$D_N = \frac{\sum n_i D_i}{\sum n_i}, \tag{5}$$

where n_i is the number of particles whose diameter is D_i; (b) the volume to surface ratio diameter D_s, where

$$D_s = \frac{\sum n_i D_i^3}{\sum n_i D_i^2}, \tag{6}$$

TABLE II. Cumulative average diameters for latex No. 235.

Groups	No. of particles	Cumulative average diameter-A D_N	D_s	D_w
I	420	482	701	863
I, II	840	523	739	860
I, II, III	1260	539	736	846
I, II, III, IV	1680	531	720	835

[9] Willson, Miller, and Rowe, J. Phys. Colloid Chem. 53, 357 (1949).

and (c) the molecular weight average diameter D_w, given by

$$D_w=\left[\frac{\sum n_i D_i^6}{\sum n_i D_i^3}\right]^{\frac{1}{3}}. \quad (7)$$

The last two averages are particularly important because they can be checked by other methods, namely, soap titration and light scattering, respectively. Intercomparisons of results obtained by these various methods will be published in the near future.

To establish the reliability of particle diameter measurements for any given sample, two pure soap Type II GR-S latices were selected for extended evaluation. These were latex No. 235, prepared using pure sodium myristate, and latex No. 236, prepared with sodium laurate. Table I shows results of the measure-measurements and calculations of particle size for these samples. In the case of the myristate latex six photographs involving several separate brominations and mounts were inspected for a total count of 1680 particles, while in the case of the laurate latex four photographs were measured for a total count of 1436 particles. For latex No. 235 prints A-1-3 and A-2-2 were combined and treated as a single picture because the number of particles on each was too small for comparison with the other prints. The various averages shown in the table for each latex indicate a satisfactory concordance from print to print, and agree well with the weighted average diameters calculated on the basis of the total particle counts and shown in the bottom horizontal column of the table.

TABLE III. Cumulative average diameters for a commercial Type III GR-S latex (NS-16).

No. of pictures	No. of particles	D_N (A)	D_s (A)	D_w (A)
1	181	380	730	1160
2	351	400	780	1150
3	526	380	750	1140
4	734	360	750	1140

However, Table I gives only a rough index of reproducibility of results, since the number of particles counted in each picture was not the same. To allow for this difference, the 1680 particles counted for latex No. 235 were divided without bias into four groups of 420 particles each, and the averages recalculated for each group. The results thus obtained gave an average deviation from the means of ±40A (±7.5 percent) in D_N, ±33A (±4.6 percent) in D_s, and ±30A (±3.6 percent) in D_w.

Table II gives the diameter averages obtained for latex No. 235 when calculated on a cumulative basis, i.e., when calculated on the basis of 420, 840, 1260, and the total of 1680 particles. From this table it is evident that, at least for this latex, a satisfactory average value of D_N could have been obtained from a count of *ca* 800–900 particles, whereas for calculation of D_s and D_w about 500 particles would have sufficed to give results within *ca* 3 percent of the final cumulative average.

For concordant results with most of the latices examined, fewer than 1000 particles had to be counted. The greatest difficulty in getting satisfactory counts was

TABLE IV. Particle size of synthetic latices.

Latex No.	Description[a]	No. of particles	Average diameter-A D_N	D_s	D_w	Type of distribution
236	Type II (72:28 BD-ST, Na laurate)	1436	400	570	730	1–1
238	Type II (72:28 BD-ST, Na laurate)	203	410	630	740	2–1
240	Type II (72:28 BD-ST, Na oleate)	590	460	690	830	2–1
235	Type II (72:28 BD-ST, Na myristate)	1680	530	720	840	1–1
220	Type II (72:28 BD-ST, commercial)	275	580	790	890	2–1
243	Type II (72:28 BD-ST, Na palmitate)	525	500	790	940	2–2
242	Type II (72:28 BD-ST, Na oleate)	439	530	1100	1180	2–1
NS-16	Type III (50:50 BD-ST, commercial)	734	360	750	1140	2–2
NS-14	Type III (50:50 BD-ST, commercial)	1707	460	790	1230	2–1
FR	Type III (50:50 BD-ST, commercial)	1714	430	790	1340	1–2
NS-11	Type III (50:50 BD-ST, commercial)	1973	500	1040	1580	2–2
NS-28	Type III (50:50 BD-ST, commercial)	242	520	1130	1620	2–2
FS-2	Type III (50:50 BD-ST, commercial)	1171	500	1050	1640	2–2
628	Polystyrene (Na oleate)	744	350	770	1310	2–2
629	Polystyrene (Na oleate)	858	400	1300	1610	2–2
630	Polystyrene (Na oleate)	472	670	1270	1620	2–2
GL-25	Type V (70:30 BD-ST, mixed soaps)	638	490	1630	2080	2–1
GL-4	Type V (70:30 BD-ST, mixed soaps)	422	550	2070	2470	2–1
GL-13	Type V (70:30 BD-ST, mixed soaps)	789	540	2340	2750	2–1
G-4	60:40 BD-ST (seeded)	318	730	2300	2450	2–1
G-5	60:40 BD-ST (seeded)	483	670	3700	3920	2–1
3–4	40:60 BD-ST	148	...	3400	3460	2–1

[a] BD = butadiene; ST = styrene.

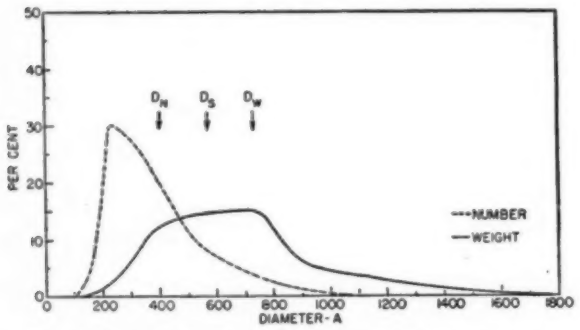


FIG. 2. Particle size distribution in latex 236. (1–1 type.)

experienced in cases in which the distribution was very broad. Since large-size particles have a tremendous influence on weight or on volume-surface area averages, the presence or absence of a few of the largest particles in any photograph modified the results greatly. When special care was taken, however, even commercial Type III GR-S latex, which generally exhibits a wide particle size distribution, gave results of satisfactory concordance, as is demonstrated in Table III. Here the cumulative averages show reproducibility that falls well within the limitations of the electron microscope technique of obtaining particle size.

PARTICLE SIZE RESULTS

Table IV summarizes the results obtained in the present study. The second column gives the monomer charge compositions of the latices, and whenever possible, the commercial type and emulsifiers employed in the latex preparation. Where specific compounds are listed, these refer to laboratory preparations made with the indicated C. P. soaps. In turn the third column lists the number of particles measured in each case while the next three columns give the corrected values of D_N, D_s, and D_w calculated from the measurements.

Comparison of the latices on the basis of the D_w averages shows the Type II latices to be of smallest size, followed in order by Type III, the polystyrene latices, Type V, and the special preparations at the bottom of the table. The largest average diameter observed here was $D_w = 3920$A for the seeded G-5 sample.

The data in Table IV reveal a considerable spread in the values of the various averages for a given sample. In every case the value of D_w is highest, D_s is intermediate, and D_N smallest. Again, in all cases but one (polystyrene latex No. 628), D_s lies closer to D_w than to D_N. The fact that the diameters calculated depend upon the method of averaging is an indication that the systems under consideration are not homogeneous in size. Were the systems homogeneous, all the averages would yield the same diameter. Furthermore, the results show that the ratios of, say, D_w to D_N are not constant for the various latices, and hence the indications are also that the distributions are not the same for the various latices.

PARTICLE SIZE DISTRIBUTION

To investigate the particle size distributions of the latices, number and weight fraction integral distribution curves for each latex were first constructed, and these were then differentiated graphically to obtain the number and weight differential distribution curves. The curves thus obtained can be classified into the four types shown in Figs. 2–5. These all differ from each other in the number of maxima. Thus the type indicated in Fig. 2 presents a single maximum in the number, as well as in weight distribution, and hence we designate this distribution as being of the 1–1 type. Again, Fig. 3 shows a single maximum in the number curve and two in the weight curve, and hence the distribution is of the 1–2 type. On this basis the distributions given in Figs. 4 and 5 are of the 2–1 and 2–2 type, respectively. Of the 22 latices investigated here two were found to be of the 1–1 type, one of the 1–2 type, eleven of the 2–1 type, and eight of the 2–2 type. The particular type of distribution exhibited by each latex sample is indicated in the last column of Table IV.

In connection with Fig. 4 it should be remarked that although the number distribution curve shows here a discontinuity, with no particles being found between *ca* 350 and 700A, such is not always the case. Thus in Fig. 5, where two maxima in number distribution also occur, the distribution of particle sizes is continuous.

Inspection of Figs. 2–5 makes it evident that different conclusions in regard to particle size distribution may be arrived at from the number and weight distribution curves for a given latex. Thus, from the number curves in Figs. 2–4, one may surmise that the first two latices have relatively narrow distributions, while that of the third is very broad. On the other hand, the weight curves for these latices would lead to just the reverse conclusions. Actually, both conclusions are valid when considered in connection with the property of the system being measured. For properties dependent on the number of particles, the number distribution is of consequence, while for properties dependent on weight, such as light scattering, the weight distribution is controlling and determines, along with the average diameter, the behavior of the system.

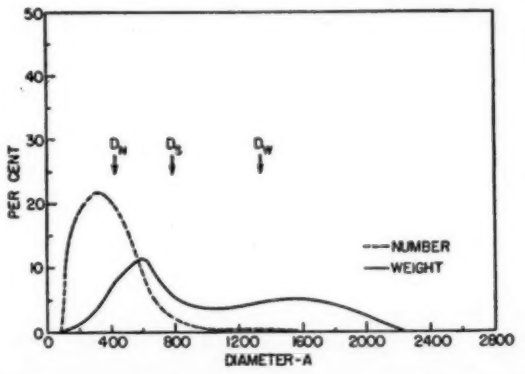


FIG. 3. Particle size distribution in latex FR. (1–2 Type.)

It is not possible to present now a detailed explanation of the nature of the distribution curves. To do so would require a careful and detailed correlational study of the relation between polymerization recipe, conditions, and mechanism on the one hand, and particle size and its distribution on the other. No such study has as yet been made.

DISCUSSION AND CONCLUSIONS

Although with proper technique electron microscopy may be a suitable and quite rapid means of determining the particle size of homogeneous systems, the method is tedious, time consuming, and full of pitfalls when applied to systems of broad particle size distribution. The difficulties encountered depend on whether the number or molecular weight average diameters are

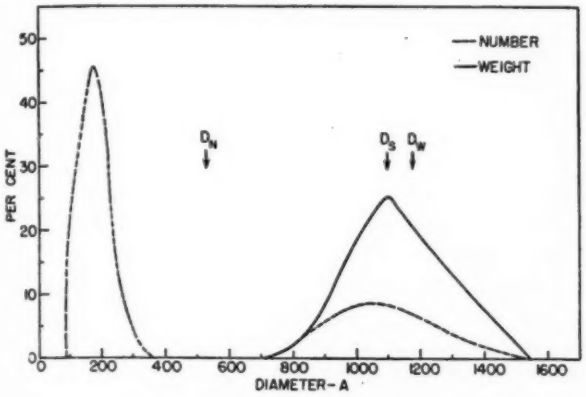


FIG. 4. Particle size distribution in latex 242. (2–1 Type.)

sought. For number average size determination it is quite evident that all the particles, small as well as large, must be counted in a proper sample of the system. However, in view of the practical limits of resolution of the microscope and the background structure of the support film it is frequently very difficult to identify and measure the small particles. Consequently, great care must be exercised in the preparation of micrographs for such counts and in the measurement of all particles which may be recognizable. In general, experience shows that when the distribution is broad more particles must be counted to obtain a reliable number average size and distribution than for the other and heavier weightings considered here. However, even with all care, it is impossible to know definitely how reliable a number count or distribution is, for there are no independent methods available for ascertaining these.

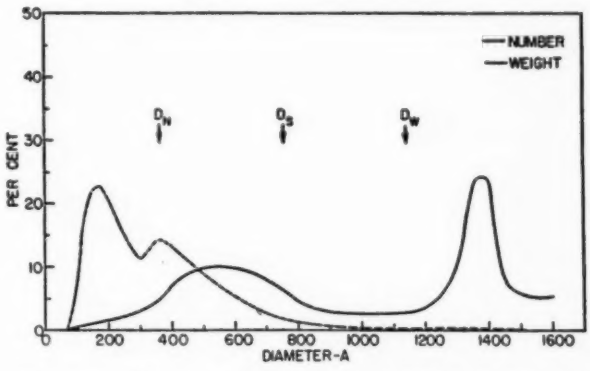


FIG. 5. Particle size distribution in latex NS-16. (2–2 Type.)

Still, electron microscopy is the only means by which the number average size and distribution of particles in systems such as latex has been determined.

In the case of weight average diameter determination the many small particles contribute very little to the total weight, and hence they may, in most instances, be practically neglected. The important point here is to obtain a true sampling of the larger particles. The nonappearance of several large particles in a micrograph may modify drastically both the average and the weight distribution obtained for a system. Consequently it is essential to make sure that the field photographed is representative, and that sufficient particles are taken to give a reliable average and distribution. We have usually found that satisfactory concordance for this type of average can be obtained with fewer particle counts than for number average. Further, in this instance the reliability of the results can be checked by a suitable light scattering method to ascertain whether the procedures employed are satisfactory.

For D_s averages the problems encountered are intermediate between the above. Both ends of a distribution are still important, but to a lesser degree than before. Here again the results can be checked independently by a soap titration method to be published in the near future.

ACKNOWLEDGMENTS

This investigation was sponsored by the Synthetic Rubber Division, Reconstruction Finance Corporation, as part of the Government Synthetic Rubber Program, and was first reported in August, 1948.

The authors are also indebted to I. N. Ulevitch and M. E. Elder of this laboratory for their assistance with a portion of this work.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 8 AUGUST, 1952

The Dirac Delta-Function and the Summation of Fourier Series

M. R. SPIEGEL
Rensselaer Polytechnic Institute, Troy, New York
(Received March 24, 1952)

The summation of Fourier series is attained by use of ordinary linear differential equations with constant coefficients having inputs that depend on certain combinations of the Dirac delta-function. The method is illustrated by several examples representative of those series that arise in applications.

I. INTRODUCTION

IN a recent article[1] L. A. Pipes demonstrated a means whereby certain Fourier series, which might arise in actual practice, could be summed. His method, based upon recognition of certain series after making use of Laplace transformation, seems to be more difficult than the method that will be proposed in the present article. There are certain definite advantages of this method over Pipes' method, as will be seen.

The method of this article is heuristic, since divergent series and impulse functions are used. The author, nevertheless, has no doubts that the procedures involved may be made mathematically rigorous. However, the author has found that in the process of rigorization, the ease and facility of the manipulations involved is lost, and the calculations may become too laborious. As a consequence the author wishes to emphasize that he will use heuristic means of arriving at the sum of the Fourier series and then prove that the sum arrived at actually is the sum corresponding to the Fourier series.

Briefly the method involves certain correspondences of the Dirac delta-function and trigonometric sums. An ordinary linear differential equation for the sum function involving only constant coefficients is obtained. The inhomogeneous portion of the differential equations involves combinations of Dirac delta-functions.[2] The differential equations can be easily solved either with or without the Laplace transform, and the required sum of the Fourier series can thus be obtained. The calculations are quite straightforward, and the method is easily applicable to Fourier series which, on the surface, appear very difficult to sum. After a few preliminary remarks in Sec. II, we shall illustrate the procedures necessary by several examples.

II. PRELIMINARY CONSIDERATIONS

The Dirac delta-function is by definition that function such that

$$\int_{-\infty}^{\infty}\delta(x-a)F(x)dx=F(a), \tag{1}$$

irrespective of the functions $F(x)$. This function $\delta(x)$ is taken to be an even function, that is, $\delta(-x)=\delta(x)$.

Given a function $f(x)$ defined in the interval $(-\pi, \pi)$ and defined by periodic extension, period 2π, outside of this interval, we associate the Fourier coefficients

$$\begin{Bmatrix} a_k \\ b_k \end{Bmatrix}=\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\begin{Bmatrix} \cos kx \\ \sin kx \end{Bmatrix}dx \tag{2}$$

and define the Fourier series corresponding to this function to be the series

$$\frac{a_0}{2}+\sum_{k=1}^{\infty}\{a_k\cos kx+b_k\sin kx\}, \tag{3}$$

where the coefficients a_k, b_k are given in Eq. (2). We shall for the purposes of the present article restrict all of our Fourier series to have period 2π and take our sum function corresponding to the Fourier series to be defined between $-\pi$ and π. If we do this, there will be no loss in generality, and our task will be greatly simplified.

Consider now the case where $f(x)$ in Eq. (2) is the Dirac delta-function $\delta(x)$. By virtue of the fact that $\delta(x)$ is an even function, we have $b_k=0$, and from Eq. (2) and the property of the delta-function defined in Eq. (1) we have

$$a_k=2/\pi\int_0^{\pi}\delta(x)\cos kx dx=2/\pi. \tag{4}$$

Hence, the Fourier cosine series corresponding to $\delta(x)$ is

$$(1/\pi)+(2/\pi)\cos x+(2/\pi)\cos 2x+(2/\pi)\cos 3x+\cdots, \tag{5}$$

and we will write formally

$$(1/\pi)+(2/\pi)\cos x+(2/\pi)\cos 2x+\cdots=\delta(x) \tag{6}$$

or

$$\tfrac{1}{2}+\sum_{n=1}^{\infty}\cos nx=(\pi/2)\delta(x). \tag{7}$$

It should be observed that, rigorously, the left-hand side of the series in Eq. (7) does not converge in the usual sense. We will therefore say that Eq. (7) holds in the heuristic sense. (The author is reminded here of the remark due to Heaviside that went something like this. "This series, according to the mathematicians, is divergent; hence, we should be able to do something useful with it." While the author is not a proponent of such a doctrine, he does believe that heurism has its

[1] L. A. Pipes, J. Appl. Phys. **21**, 298 (1950).
[2] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Clarendon Press, Oxford), second edition, pp. 72–78.

place in securing results—after which they should be verified according to the rules of rigorous mathematics.)

If we replace x by $2x$ in Eq. (7), we obtain

$$\tfrac{1}{2}+\sum_{n=1}^{\infty}\cos 2nx=(\pi/2)\delta(2x), \tag{8}$$

and subtraction of Eqs. (7) and (8) yields the result

$$\cos x+\cos 3x+\cos 5x+\cdots=(\pi/2)[\delta(x)-\delta(2x)]. \tag{9}$$

It should be pointed out that one might also have obtained the relation

$$\cos x+\cos 3x+\cos 5x+\cdots=(\pi/4)[\delta(x)-\delta(\pi-x)] \tag{10}$$

by using Eq. (7) with x replaced by $\pi-x$.

Several results similar to Eqs. (7), (9), and (10) could easily be obtained, but it turns out that for most of the series for which summation may be obtained Eq. (7), Eq. (8), or Eq. (9) is all one has to know. The use of these series and the general procedure involved will be indicated in the next section.

III. ILLUSTRATIVE EXAMPLES

In this section we shall find the sums of some Fourier series that the author believes are sufficiently representative of those series that normally arise in practice, and that can be summed.

Example 1

Find the function whose Fourier series is

$$\frac{1}{\pi}+\frac{\sin t}{2}-\frac{2}{\pi}\left[\frac{\cos 2t}{3}+\frac{\cos 4t}{15}+\frac{\cos 6t}{35}+\cdots\right]. \tag{11}$$

Let

$$Y(t)=\frac{\cos 2t}{3}+\frac{\cos 4t}{15}+\frac{\cos 6t}{35}+\cdots. \tag{12}$$

Then

$$Y'(t)=\frac{-2\sin 2t}{3}-\frac{4\sin 4t}{15}-\frac{6\sin 6t}{35}-\cdots, \tag{13}$$

and

$$Y''(t)=-\frac{4\cos 2t}{3}-\frac{16\cos 4t}{15}-\frac{36\cos 6t}{35}-\cdots.$$

Thus,

$$\begin{aligned}Y''(t)+Y(t)&=-\cos 2t-\cos 4t-\cos 6t-\cdots\\&=\tfrac{1}{2}-\{\tfrac{1}{2}+\cos 2t+\cos 4t+\cdots\}\\&=\tfrac{1}{2}-(\pi/2)\delta(2t)\end{aligned}$$

by making use of Eq. (8).

Taking the Laplace transform of

$$Y''(t)+Y(t)=\tfrac{1}{2}-(\pi/2)\delta(2t), \tag{14}$$

we obtain

$$s^2y(s)-sY(0)-Y'(0)+y(s)=(1/2s)-(\pi/4), \tag{15}$$

since

$$\int_0^{\infty}e^{-st}\delta(2t)dt=\frac{1}{2}\int_0^{\infty}e^{-st/2}\delta(t)dt=\tfrac{1}{2}.$$

Also from Eq. (12) it is clear that $Y'(0)=0$. Thus, Eq. (15) gives

$$y(s)=\frac{sY(0)}{s^2+1}+\frac{1}{2s(s^2+1)}-\frac{\pi}{4(s^2+1)}. \tag{16}$$

Taking the inverse Laplace transform of Eq. (16) we have

$$Y(t)=(Y(0)-\tfrac{1}{2})\cos t+\tfrac{1}{2}-(\pi/4)\sin t. \tag{17}$$

Now Eq. (12) is an even function since $Y(-t)=Y(t)$; yet the right-hand side of Eq. (17) is not an even function because of the presence of $\sin t$. However, since the solution (17) was obtained for $t>0$, it is easy to make it an even function by writing

$$Y(t)=(Y(0)-\tfrac{1}{2})\cos t+\tfrac{1}{2}-(\pi/4)|\sin t|. \tag{18}$$

To determine the constant $Y(0)$, multiply Eq. (12) by $\cos t$ and integrate from $-\pi$ to π. By orthogonality we have

$$\int_{-\pi}^{\pi}Y(t)\cos t\,dt=0. \tag{19}$$

Now multiplying Eq. (18) by $\cos t$ and integrating from $-\pi$ to π we find

$$\int_{-\pi}^{\pi}Y(t)\cos t\,dt=\pi[Y(0)-\tfrac{1}{2}]. \tag{20}$$

From Eqs. (19) and (20) it is thus clear that $Y(0)=\frac{1}{2}$. Hence Eq. (18) becomes

$$Y(t)=\tfrac{1}{2}-(\pi/4)|\sin t|,$$

and the required sum function (11) is

$$\begin{aligned}F(t)&=(1/\pi)+(\sin t/2)-(2/\pi)Y(t)\\&=\sin t/2+(|\sin t|/2)\end{aligned} \tag{21}$$

or

$$F(t)=\begin{cases}0 & -\pi<t<0\\ \sin t & 0<t<\pi\end{cases}. \tag{22}$$

An expansion of the function (22) in a Fourier series yields the series (11) showing that we have indeed arrived at the correct function.

Example 2

Find the sum of the series

$$Y(x)=\sum_{n=0}^{\infty}\cos nx/n^2+1. \tag{23}$$

Differentiation yields

$$Y'(x)=\sum_{n=0}^{\infty}-n\sin nx/n^2+1, \tag{24}$$

$$Y''(x)=\sum_{n=0}^{\infty}-n^2\cos nx/n^2+1. \tag{25}$$

Hence,

$$Y(x)-Y''(x)=\sum_{n=0}^{\infty}\cos nx=\tfrac{1}{2}+(\pi/2)\delta(x),$$

by making use of Eq. (7). Thus,

$$Y''(x)-Y(x)=-\tfrac{1}{2}-(\pi/2)\delta(x), \tag{26}$$

and taking the Laplace transform we obtain

$$s^2y(s)-sY(0)-Y'(0)-y(s)=-(1/2s)-(\pi/2).$$

Since $Y'(0)=0$, we obtain

$$y(s)=\frac{sY(0)}{s^2-1}-\frac{1}{2s(s^2-1)}-\frac{\pi}{2(s^2-1)}.$$

Thus,

$$Y(x)=(Y(0)-\tfrac{1}{2})\cosh x-(\pi/2)\sinh x \tag{27}$$

upon inversion.

Since Eq. (23) is an even function we must write Eq. (27) as

$$Y(x)=(Y(0)-\tfrac{1}{2})\cosh x-(\pi/2)|\sinh x|, \tag{28}$$

which is now an even function as required.

To determine $Y(0)$, we integrate Eq. (23) with respect to x from $-\pi$ to π and obtain

$$\int_{-\pi}^{\pi}Y(x)dx=0. \tag{29}$$

From Eq. (28) we obtain upon integration from $-\pi$ to π

$$\int_{-\pi}^{\pi}Y(x)dx$$

$$=(Y(0)-\tfrac{1}{2})\int_{-\pi}^{\pi}\cosh x dx-(\pi/2)\int_{-\pi}^{\pi}|\sinh x|\,dx$$

$$=2(Y(0)-\tfrac{1}{2})\sinh\pi+\pi(1-\cosh\pi). \tag{30}$$

From Eqs. (29) and (30) we find

$$Y(0)=\tfrac{1}{2}+\frac{\pi}{2}\left(\frac{\cosh\pi-1}{\sinh\pi}\right).$$

Hence, Eq. (28) yields the result

$$Y(x)=\frac{\pi}{2}\left(\frac{\cosh\pi-1}{\sinh\pi}\right)\cosh x-\frac{\pi}{2}|\sinh x|, \tag{31}$$

which is the sum of the series (23).

Example 3

Find the sum of the series

$$f(t)=\frac{\pi t}{2}+\frac{4}{\pi}\left[\frac{\sin t}{1^3}+\frac{\sin 3t}{3^3}+\frac{\sin 5t}{5^3}+\cdots\right].$$

Let

$$Y(t)=\frac{\sin t}{1^3}+\frac{\sin 3t}{3^3}+\frac{\sin 5t}{5^3}+\cdots. \tag{32}$$

Then

$$Y'(t)=\frac{\cos t}{1^2}+\frac{\cos 3t}{3^2}+\frac{\cos 5t}{5^2}+\cdots,$$

$$Y''(t)=-\frac{\sin t}{1}-\frac{\sin 3t}{3}-\frac{\sin 5t}{5}-\cdots, \tag{33}$$

and

$$Y'''(t)=-\cos t-\cos 3t-\cos 5t-\cdots.$$

Making use of Eq. (9) we have

$$Y'''(t)=-(\pi/2)[\delta(t)-\delta(2t)]. \tag{34}$$

Hence, by Laplace transformation we obtain

$$s^3y(s)-s^2Y(0)-sY'(0)-Y''(0)$$
$$=-(\pi/2)[1-\tfrac{1}{2}]=-(\pi/4),$$

so that since $Y(0)=Y''(0)=0$, we have

$$y(s)=(Y'(0)/s^2)-(\pi/4s^3).$$

Inversion of the Laplace transform yields

$$Y(t)=tY'(0)-(\pi/8)t^2. \tag{35}$$

Since $Y(t)$ must be an odd functions of t, we will have

$$Y(t)=\begin{cases}tY'(0)-(\pi/8)t^2 & 0<t<\pi\\ tY'(0)+(\pi/8)t^2 & -\pi<t<0\end{cases}. \tag{36}$$

Multiplying Eq. (32) by $\sin t$ and integrating from $-\pi$ to π, we obtain

$$\int_{-\pi}^{\pi}Y(t)\sin t dt=\pi. \tag{37}$$

From Eq. (36) we obtain

$$\int_{-\pi}^{\pi}Y(t)\sin t dt=2\int_{0}^{\pi}(tY'(0)-(\pi/8)t^2)\sin t dt$$
$$=2\pi Y'(0)+\pi-(\pi^3/4). \tag{38}$$

From Eqs. (37) and (38) it is thus clear that $Y'(0)=\pi^2/8$. Hence Eq. (36) becomes

$$Y(t)=\begin{cases}(\pi t/8)(\pi-t) & 0<t<\pi\\ (\pi t/8)(\pi+t) & -\pi<t<0,\end{cases}$$

and so the required sum of the series $f(t)$ is

$$f(t)=\begin{cases}\pi t-t^2/2 & 0<t<\pi\\ \pi t+t^2/2 & -\pi<t<0\end{cases}. \tag{39}$$

As a last example we choose one that has already been considered in the paper by Pipes.

Example 4

Find the sum of the series

$$Y(t)=4\left[\frac{\sin t}{1}+\frac{\sin 3t}{3}+\frac{\sin 5t}{5}+\cdots\right]. \tag{40}$$

We have

$$Y'(t)=4[\cos t+\cos 3t+\cos 5t+\cdots]=2\pi[\delta(t)-\delta(2t)].$$

Taking the Laplace transform, we find

$$sy(s)-Y(0)=2\pi[1-\tfrac{1}{2}]=\pi, \tag{41}$$

from which we obtain, since $Y(0)=0$,

$$y(s)=\pi/s.$$

By inverting, we find

$$Y(t)=\pi.$$

Since $Y(t)$ is odd, we have

$$Y(t)=\begin{cases}\pi & 0<t<\pi\\ -\pi & -\pi<t<0\end{cases}, \tag{42}$$

which yields the desired sum. The graph for this function may, of course, easily be obtained from Eq. (42) for the interval $(-\pi, \pi)$, and may be extended periodically, period 2π, outside of this interval.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 8 AUGUST, 1952

Electromagnetic Back-Scattering from Cylindrical Wires*

C. T. Tai
Stanford Research Institute, Stanford, California

(Received April 10, 1952)

The problem of electromagnetic back-scattering or radar response from cylindrical wires has been investigated using a variational method. The relation between this method and the induced emf method is discussed. To demonstrate the flexibility of the variational calculation different trial functions have been used to determine the numerical values of the back-scattering cross section for the case of broadside incidence. The boundary condition regarding the currents at the ends of the wire is also carefully examined.

I. INTRODUCTION

THE theoretical problem of electromagnetic back-scattering from perfectly conducting cylindrical wires has been investigated by a number of writers. Previous work on this problem is contained in an article by Van Vleck, Bloch, and Hamermesh[1] in which the unpublished work of Chu is also discussed. This problem represents, of course, just one aspect of the general theory of a receiving antenna, which was previously studied by Hallén[2] and King and Harrison.[3] In the work of Van Vleck *et al.*, the calculation of the back-scattering cross section of the wires was based upon two "independent methods." The first method, which will be designated as the induced emf method, was first proposed by Chu to investigate the response of resonant wires and long wires. In essence, it is an alternative version of the classical emf method used by many earlier investigators in studying the properties of transmitting antennas. The second method, which will be designated as an integral equation method, is similar to the original method contributed by Hallén,[2] and the modified ones of Gray[4] and King and Middleton.[5] However, no iteration is used. Because of the different types of approximations involved in the two methods, and hence the numerical results, it is difficult to conclude which method gives a better over-all answer to the problem.

The main purpose of this paper is to present an independent calculation of the back-scattering cross section of the wires based upon a variational method which was originally due to Schwinger.[6] First of all, this formulation eliminates the arbitrariness in choosing the thickness factor which has occurred in previous methods. Secondly, as a result of the stationary characteristics of the variational expression, it is possible to determine rather accurately the value of the back-scattering cross section by using appropriate trial functions for the current distribution on the wire or by finding the higher order solutions. Finally, in the present formulation the role played by the "boundary condition," that the current vanishes at the ends of the wire, can be examined quantitatively. Certain ambiguities occurring in previous work on the general problem of receiving antennas can thus be clarified in the light of this analysis.

* The investigation of this problem was supported by Air Force Contract AF 19(122)78.

[1] Van Vleck, Bloch, and Hamermesh, J. Appl. Phys. **18**, 274 (1947).

[2] E. Hallén, Nova Acta, (Upsala) **11**, 1 (1938).

[3] R. King and C. W. Harrison, Jr., Proc. Inst. Radio Engrs. **32**, 18 (1944); R. King, "Graphical Representation of the Characteristics of Cylindrical Antennas," Cruft Lab. Tech. Rept. No. 20 (Harvard University, Cambridge, Massachusetts, 1947).

[4] M. C. Gray, J. Appl. Phys. **15**, 61 (1944).

[5] R. King and D. Middleton, Quart. Appl. Math. **3**, 302 (1946).

[6] J. Schwinger, Phys. Rev. **72**, 742 (1947).

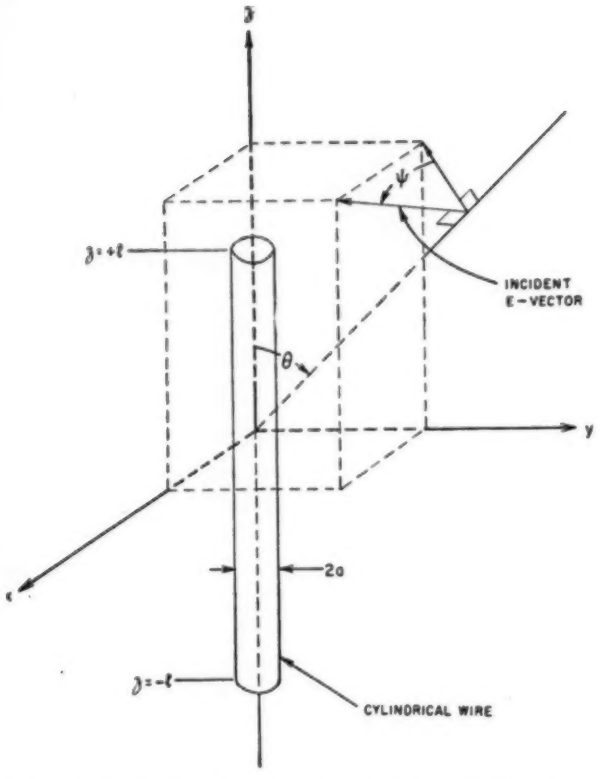


FIG. 1. Orientation of the incident E-vector and the position of the cylindrical wire.

II. THE INTEGRAL EQUATION AND THE VARIATIONAL EXPRESSION FOR THE BACK-SCATTERING CROSS SECTION

It is known in antenna theory[2,3] that the integro-differential equation which governs the current distribution on a perfectly conducting, thin, cylindrical wire excited by a plane electromagnetic wave (Fig. 1) is given to a good approximation by

$$\frac{d^2A_z}{dz^2}+k^2A_z=-\frac{k^2}{j\omega}E_0\sin\theta\cos\psi e^{jkz\cos\theta}, \tag{1}$$

with

$$A_z=\frac{\mu}{4\pi}\int_{-l}^{l}I(z')\frac{e^{-jkR}}{R}dz', \tag{2}$$

where the following notation (in the MKS system of units) is used:

E_0 = amplitude of the electric field of the incident wave,
$k=2\pi/\lambda$,
λ = free space wavelength,
$I(z')$ = current in the wire at z',
$R=[(z-z')^2+a^2]^{\frac{1}{2}}$,
a = radius of wire,
$2l$ = length of wire,
θ = angle between the wire and the wave normal, and
ψ = polarization angle shown in Fig. 1.

Equations (1) and (2) may be combined to form a simple integral equation of the form

$$E_0\sin\theta\cos\psi e^{jkz\cos\theta}=\frac{j\eta k}{4\pi}\int_{-l}^{l}I(z')\left(1+\frac{\partial^2}{k^2\partial z^2}\right)\frac{e^{-jkR}}{R}dz', \tag{3}$$

where

$$\eta=\left(\frac{\mu}{\epsilon}\right)^{\frac{1}{2}}=120\pi \text{ ohms}.$$

Equation (3) is equivalent to the original integral equation derived by Hallén[2] and used later by many other writers.[1,3] Hallén's equation is obtained by solving Eq. (1) for A_z and then substituting the result into Eq. (2), giving

$$\frac{E_0\cos\psi e^{jkz\cos\theta}}{\sin\theta}+C_1\sin kz+C_2\cos kz=\frac{j\eta k}{4\pi}\int_{-l}^{l}I(z')\frac{e^{-jkR}}{R}dz', \tag{4}$$

where C_1 and C_2 are two arbitrary constants resulting from the general solution for the homogeneous differential equation contained in Eq. (1). According to Hallén's analysis, these two constants are ultimately determined by imposing the "boundary condition" that the current be vanishing at the ends of the wire. It will be shown later that this condition is really not a necessary one for obtaining the solution of the problem, and that it can be relaxed in the variational formulation which is based upon Eq. (3).

To obtain the variational expression for the back-scattering cross section, a quantity which is stationary with respect to a small variation of $I(z)$ is extracted from the integral equation. By multiplying Eq. (3) by $I(z)$ and integrating over the range $z=\pm l$, and then dividing by

$$\left[k\sin\theta\int_{-l}^{l}I(z)e^{jkz\cos\theta}dz\right]^2,$$

the following equation is obtained:

$$\frac{1}{S}=\frac{j\int_{-l}^{l}\int_{-l}^{l}I(z)I(z')K(z-z')dzdz'}{\left[k\sin\theta\int_{-l}^{l}I(z)e^{jkz\cos\theta}dz\right]^2}, \tag{5}$$

where

$$S=\frac{\eta k^2\sin\theta\int_{-l}^{l}I(z)e^{jkz\cos\theta}dz}{4\pi E_0\cos\psi}, \tag{6}$$

$$K(z-z')=k\left(1+\frac{\partial^2}{k^2\partial z^2}\right)\frac{e^{-jkR}}{R}. \tag{7}$$

Considering S as a parameter, it can be easily verified that the expression (5) is stationary with respect to first-order variations of $I(z)$ about its correct value determined by Eq. (3). According to definition, the back-scattering cross section $\sigma(\theta, \psi)$ is defined as

$$\sigma(\theta, \psi)=\frac{4\pi R_0^2|E_\theta \cos\psi|^2}{E_0^2}, \tag{8}$$

where E_θ denotes the field produced by the induced current $I(z)$ at a large distance R_0 in the direction opposite that of the incident wave; that is,

$$E_\theta=\frac{j\eta k \sin\theta e^{-jkR_0}}{4\pi R_0}\int_{-l}^{l} I(z)e^{jkz\cos\theta}dz. \tag{9}$$

When the common integral contained in Eqs. (6) and (9) is eliminated, and the resultant expression for E_θ is substituted into Eq. (8),

$$\sigma(\theta, \psi)=\frac{4\pi \cos^4\psi}{k^2}|S|^2. \tag{10}$$

Equation (10) is the expression relating σ and S. The latter which is stationary with respect to small variation of $I(z)$ is to be evaluated using the expression given by Eq. (5). In obtaining the numerical values of $\sigma(\theta, \psi)$, it is convenient to eliminate the ψ-dependence by introducing the average value of $\sigma(\theta, \psi)$ over all values of ψ. Denoting the latter by $\sigma(\theta)$, we have

$$\sigma(\theta)=\frac{1}{\pi}\int_0^\pi \sigma(\theta, \psi)d\psi=\frac{3\pi}{2k^2}|S|^2, \tag{11}$$

or

$$\sigma(\theta)/\lambda^2=3|S|^2/8\pi. \tag{12}$$

This concludes the derivation of the variational expression for $\sigma(\theta)$ as given by Eqs. (5) and (12).

Before any numerical computation is made, it is of interest to examine the difference between the variational method and the emf method. The essential expression used in the emf method as given by Van Vleck *et al.*, however, will be presented first in a rather different manner in order to show its resemblance to the variational method.

Consider the integral equation given by Eq. (3). If this equation is multiplied by $I^*(z)$, the complex conjugate of $I(z)$, and integrated over the range $z=\pm l$, the following identity is obtained from the real part of the resultant equation:

$$Re\left[E_0 \sin\theta \cos\psi \int_{-l}^{l} I^*(z)e^{jkz\cos\theta}dz\right]$$

$$=Re\left[\frac{j\eta}{4\pi}\int_{-l}^{l}\int_{-l}^{l} I^*(z)I(z)K(z-z')dzdz'\right], \tag{13}$$

where $K(z-z')$ is defined by Eq. (7). Equation (13) is simply a mathematical statement of the law of conservation of energy. The term on the left represents the power due to the induced emf, and the right-hand term is equal to the reradiated power due to the induced current. In the paper by Van Vleck *et al.*, the latter is replaced by an equivalent integral of the Poynting vector evaluated at a large sphere around the wire. Equation (13) is the essential expression used in the emf method. It provides a relationship between the magnitude of the induced current and the amplitude of the incident field. By assuming a certain current distribution along the wire, it is possible to calculate the back-scattering cross section using Eqs. (8), (9), and (13). The result, however, depends rather critically upon the specific form of the current distribution function which was assumed in the calculation. This will be demonstrated later when comparison of the numerical results obtained from the different methods is made. It is perhaps worth mentioning that according to the presentation described here, the emf method is actually not entirely independent of the integral equation method in view of the fact that the essential equation of the emf method as given by Eq. (13) is derived from the fundamental integral equation, although the imaginary part of the equation is not considered.

III. OBLIQUE INCIDENCE

As far as the numerical result is concerned, the accuracy of the variational solution usually depends upon the order of solution that one attempts to obtain. For convenience, a zeroth-order solution is designated as the one which is obtained from Eq. (5) by assuming $I(z)=I_0f_0(z)$ where I_0 is an arbitrary constant and $f_0(z)$ is a given trial function. Since the expression for S is homogeneous with respect to $I(z)$, the zeroth-order solution for S, and hence $\sigma(\theta)/\lambda^2$ is independent of the arbitrary constant I_0. A first-order solution is obtained by assuming $I(z)=I_0[f_0(z)+Af_1(z)]$, where A is an additional arbitrary constant and $f_0(z)$ and $f_1(z)$ are two independent given functions. Because of the stationary characteristics of S^{-1} with respect to $I(z)$, the arbitrary constant A is determined by putting $\partial/\partial A(1/S)=0$. Higher order solutions can be constructed by using more terms in the trial function.

Before various types of trial functions are proposed, the behavior of the current at the ends of the wire must be examined more critically. Since the publications of Hallén's work, it has generally been accepted that in order to determine the arbitrary constants C_1 and C_2 contained in Eq. (4), a "boundary condition" must be imposed upon the solution that the current be vanishing at the ends, i.e.,

$$I(\pm l)=0. \tag{14}$$

From the physical point of view, there is good reason to believe that this condition is acceptable though it does not represent the exact situation. One may recall, for example, that in the case of a transmitting antenna, it

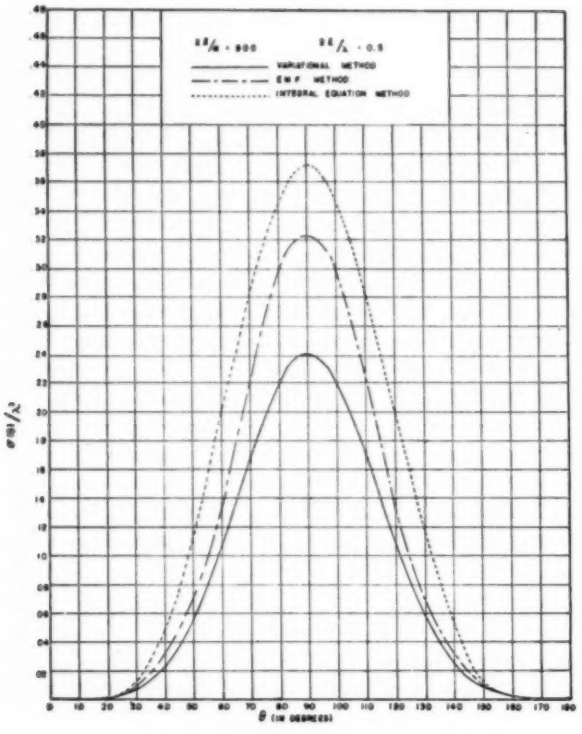


FIG. 2. Angular distribution of response of wire.

has been shown[7] that the current at the end of the lateral surface of a thin biconical antenna is actually not identically zero but is relatively small. Similar conditions must exist at the ends of a cylindrical antenna. Since the currents at the ends of the antenna are not identically zero, Eq. (14) should not be considered as a necessary boundary condition for the problem. The relaxation of this "boundary condition" is particularly significant in the variational formulation because Eq. (14) is not an indispensible condition which the variational solution has to fulfill. The arbitrary constants appearing in Hallén's integral equation, Eq. (4), do not exist in this formulation. As a matter of fact, in choosing the trial function for the variational solution, one more degree of freedom is gained if this "boundary condition" is not imposed upon the solution.

Consider first the case of oblique incidence. It is quite obvious by referring to previous work[1-3] that the simplest type of function that will adequately describe the current distribution along the wire is of the following form:

$$I(z)=I_0'[e^{jkz\cos\theta}+\alpha\cos kz+\beta\sin kz]\cdots, \quad (15)$$

where I_0' α, and β are three arbitrary constants. If no restriction is placed upon the currents at the ends, as discussed previously, the solution for S is obtained by substituting $I(z)$ into Eq. (5). The result will represent a second-order solution with the arbitrary constants α and β determined by putting $\partial/\partial\alpha(1/S)=0$ and $\partial/\partial\beta(1/S)=0$. On the other hand, if the "boundary condition" as defined by Eq. (14) is applied to Eq. (15), then the expression for $I(z)$ reduces to

$$I(z)=I_0\{[\cos kz\cos(kl\cos\theta)-\cos kl\cos(kz\cos\theta)] +A[\sin kz\sin(kl\cos\theta)-\sin kl\sin(kz\cos\theta)]\}, \quad (16)$$

where I_0 and A are two new arbitrary constants. If expression (16) is used as a trial function for $I(z)$ in computing S, the solution will represent a first-order solution with the arbitrary constant A determined by putting $\partial/\partial A(1/S)=0$. It is recalled that in the paper by Van Vleck *et al.*, an expression for the current similar to that given by Eq. (16) was used in the EMF method. The constants I_0 and A, however, were partly fixed in an arbitrary manner. As a result, the value of $\sigma(\theta)/\lambda^2$ calculated by Van Vleck *et al.* depends very much upon the particular choice of the expansion parameter Ω, as will be shown later.

Because of the complicated algebraic equations that would be involved in the second-order solution, the numerical results for the case of oblique incidence are based upon a first-order solution. The effect of relaxing the "boundary condition" will be discussed in more detail in the next section, which deals with the broadside incidence. By substituting the expression for $I(z)$ given by Eq. (16) into Eq. (5), and evaluating all the integrals, the following expression for S^{-1} is obtained:

$$1/S=(\gamma_c+A^2\gamma_s)/(g_c+Ag_s)^2, \quad (17)$$

where g_c, g_s, γ_c, and γ_s denote four known parameters resulting from various integrations. The formulas for these parameters are given in Section A of the Appendix. Because of the stationary characteristics of S^{-1} with respect to $I(z)$ the constant A is determined by putting

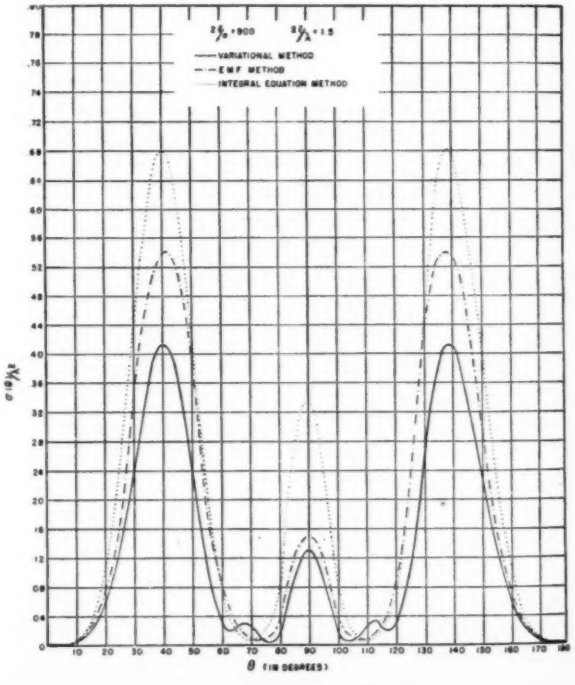


FIG. 3. Angular distribution of response of wire.

[7] C. T. Tai, J. Appl. Phys. **20**, 717 (1949).

$\partial/\partial A(1/S)=0$, yielding

$$A=g_s\gamma_c/g_c\gamma_s. \tag{18}$$

The final expression for S is then given by

$$S=(g_c^2/\gamma_c)+(g_s^2/\gamma_s); \tag{19}$$

hence,

$$\frac{\sigma(\theta)}{\lambda^2}=\frac{3}{8\pi}\left|\frac{g_c^2}{\gamma_c}+\frac{g_s^2}{\gamma_s}\right|^2. \tag{20}$$

The angular distribution of the response curve based upon Eq. (20) is plotted on Figs. 2–4 for three fixed values of kl. The corresponding curves obtained by Van Vleck *et al.* using the emf method and the integral equation method are also reproduced there. In comparing these results, it appears that the variational method, in general, predicts a response somewhat smaller than the others. As far as the over-all amplitude of the response is concerned, the result, due to the integral equation method of Van Vleck *et al.*, departs more from the present solution than does the emf method. It is quite significant to observe that for the case $kl=2\pi$, the present response curve does not contain the nulls which occur in the results of the other two methods.

Before the case of normal incidence is discussed, the asymptotic expression for (20) as applied to long non-resonant wires should be examined. Without going into detail, it can be easily verified that when kl is very large compared to unity but not near $n\pi/2$ with $n=1, 2$, etc., the following asymptotic expression of Eq. (20) is valid:

$$\frac{\sigma(\theta)}{\lambda^2}=\frac{3}{8\pi}\cdot\frac{1}{[\pi/2]^2+\{\ln[\gamma ka(1-q^2)^{\frac{1}{2}}]/2\}^2}\cdot\frac{1}{(1-q^2)^2}\left[\frac{1+q^2}{2q}\sin 2qx-\frac{1-\cos 2x\cos 2qx}{\sin 2x}\right]^2, \tag{21}$$

where $q=\cos\theta$, $x=kl$, $\ln\gamma=$Euler's constant$=0.5772$. The corresponding formula derived by Chu as quoted in the paper by Van Vleck *et al.* is

$$\frac{\sigma(\theta)}{\lambda^2}=\frac{3}{8\pi}\cdot\frac{1}{[\pi/2]^2+\{\ln[\gamma ka(1-q^2)^{\frac{1}{2}}]/2\}^2}\cdot\frac{(1-q^2)\sin^2 2qx}{4q^2}. \tag{22}$$

In spite of the difference of the fine structure between Eqs. (21) and (22), both contain the many lobes characteristic of long antennas. According to Van Vleck *et al.*, the result of their emf method yields a sharp broadside lobe, i.e., $\sigma(\theta)/\lambda^2\to 0$ for $\theta\neq\pi/2$. From a consideration of the reciprocal theorem and the known properties of long transmitting antennas, their solution seems to be inadequate. When $\theta=\pi/2$, Eqs. (21) and (22) give practically the same answer, namely,

$$\frac{\sigma_\perp}{\lambda^2}=\frac{3}{8\pi}\cdot\frac{x^2}{(\pi/2)^2+(\ln\gamma ka/2)^2}. \tag{23}$$

The last expression also agrees with Van Vleck *et al.* for this special case. As will be shown later, the agreement is due to the proper choice of the thickness parameter which these authors have used in the emf method.

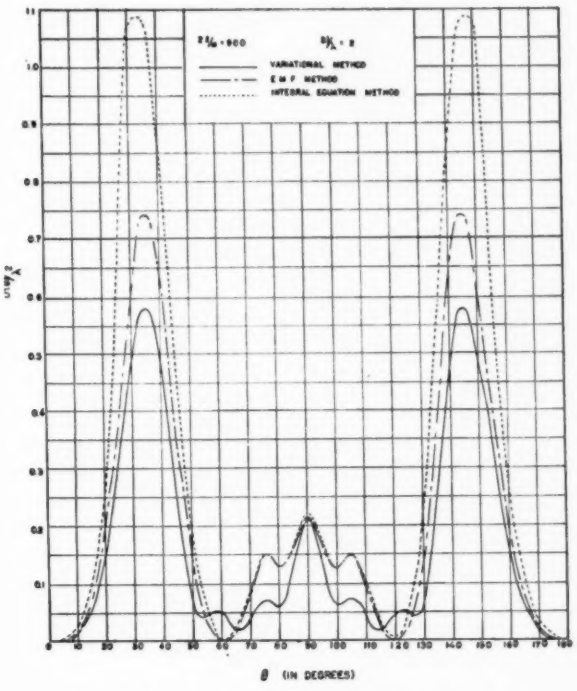


FIG. 4. Angular distribution of response of wire.

IV. BROADSIDE INCIDENCE

Although the problem of broadside incidence is only a special case of the general problem discussed in the previous section, it provides an excellent model which can be treated in great detail by the variational method. Consider first the general trial function defined by Eq. (16). When $\theta=\pi/2$, the odd part of the function disappears. The original first-order trial function, therefore, degenerates into a zeroth-order one, which becomes

$$I(z)=I_0(\cos kz-\cos kl). \tag{24}$$

The corresponding zeroth-order solution for S is then given by

$$1/S=\gamma_0/g_0^2, \tag{25}$$

and, hence,

$$\frac{\sigma_\perp}{\lambda^2}=\frac{3}{8\pi}\left|\frac{g_0^2}{\gamma_0}\right|^2, \tag{26}$$

where $\sigma_\perp$ denotes the back-scattering cross section for broadside indidence. The parameters γ_0 and g_0 are obtained by putting $\theta=\pi/2$ in the expressions for γ_c and g_c. By using the formulas for γ_0 and g_0 as given in B of

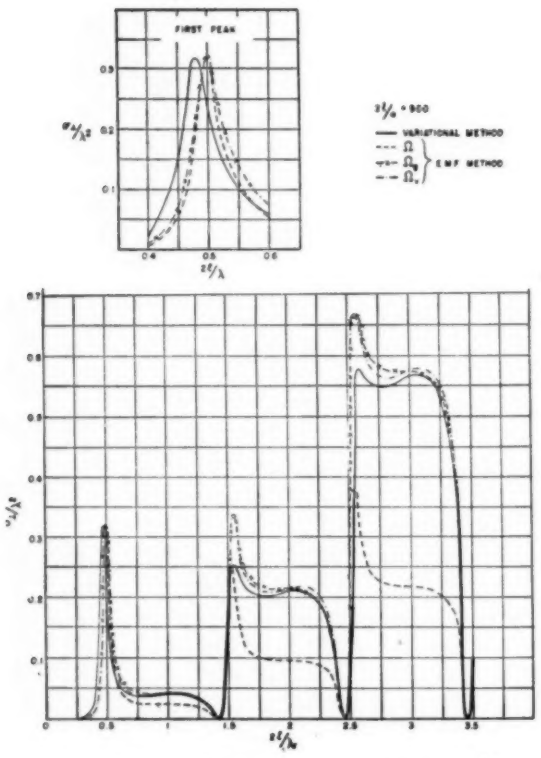


FIG. 5. Broadside response curve.

the Appendix, Eq. (26) can be written as

$$\frac{\sigma_\perp}{\lambda^2}=\frac{3}{2\pi}(x\cos x-\sin x)^2\cdot\left\{\left[(\Omega+\ln 4-2\overline{Ci}2x)\cos x+\frac{Si4x-2\cos^2 x\sin 2x}{2(x\cos x-\sin x)}\right]^2+\left[2Si2x\cos x+\frac{\overline{Ci}4x-\sin^2 2x}{2(x\cos x-\sin x)}\right]^2\right\}^{-1},\quad(27)$$

where $\Omega=2\ln 2l/a$, $x=kl$. It is recalled that the corresponding formula obtained by Van Vleck *et al.* using the emf method with the same trial function is

$$\left(\frac{\sigma_\perp}{\lambda^2}\right)_{VBH}=\frac{3}{2\pi}\times\frac{(x\cos x-\sin x)^2}{[\Omega_v\cos x]^2+\left[2Si2x\cos x+\frac{\overline{Ci}4x-\sin^2 2x}{2(x\cos x-\sin x)}\right]^2},\quad(28)$$

where Ω_v is an expansion parameter defined as follows:

$$\Omega_v=\Omega+\ln 4-2\overline{Ci}2x.\quad(29)\dagger$$

This parameter differs slightly from the one used by Gray[5] in her work dealing with the modification of Hallén's solution of the antenna problem which is given by

$$\Omega_g=\Omega+\ln 4-2\overline{Ci}2x-(\sin 2x/x).\quad(30)$$

Aside from the question which arises in regard to the choice of the expansion parameter, the emf method does not provide the information concerning the so-called resonance shift of the response curve near the first peak. Numerical calculations based upon Eqs. (27) and (28) are plotted in Fig. 5. Two more curves have been added to that figure to show the changes that would result if the expansion parameter Ω_v used by Van Vleck *et al.* were replaced by Ω or Ω_g.

Consider now the data shown in Fig. 5. It is quite obvious that one main defect that applies to the zeroth-order variational solution as well as to the emf method using the trial function of the form (24) is the vanishing of σ for certain particular values of kl which correspond to the roots of the transcendental equation, $\tan kl=kl$. To improve the solution near and at these points, a better trial function must be constructed. To show the flexibility of the variational method, and also to check the previous calculation, a first-order solution will be discussed using a set of functions which has no apparent relation to the approximate solution of the integration equation in the previous case. The new trial function is made of two harmonic functions as follows:

$$I(z)=I_0[\cos\lambda_1 z+A\cos\lambda_3 z],\quad(31)$$

where

$$\lambda_1=\pi/2l,\quad\lambda_3=3\pi/2l.$$

The constant A is determined in the same manner as explained previously. Numerical values based upon this new trial function are plotted in Fig. 6 for the range of kl from 0 to 2π. The improvement over the previous zeroth-order solution in the neighborhood of the first null is obvious, while the response curve remains practically unchanged for the range of kl between 0 and 1.2π.

V. RELAXATION OF THE "BOUNDARY CONDITION"

The last phase of this work will contain a quantitative examination of the role played by the "boundary condition," Eq. (14), in the solution of this problem. It was mentioned previously that in deriving Eq. (3) there was actually no restriction upon the current such that it should vanish at the ends of the lateral surface of the wire. Thus, if Eq. (15), which does not automatically satisfy the condition specified by Eq. (14), is used as a trial function, then for the case of broadside incidence it reduces to

$$I(z)=I_0'[1+\alpha\cos kz+\beta\sin kz].\quad(32)$$

Since no odd current is excited for this case, β must be equal to zero. The final form of the trial function can therefore be written as

$$I(z)=I_0[\cos kz-K].\quad(33)$$

† The original notation for Ω_v in the paper by Van Vleck *et al.* was Ω, which is used in this paper to denote Hallén's parameter.

The constant K is determined by substituting Eq. (33) into Eq. (5) and put $\partial/\partial K(1/S)=0$. The result based upon this trial function is plotted in Fig. 6. For $x<\pi$, the result is practically the same as the solutions discussed in the preceding section. In addition to the improvement of the shape of the response curve near the null which was present in the zeroth-order solution, the general agreement between the two first-order solutions using Eqs. (31) and (33) as two independent trial functions seems to be quite satisfactory. For long antennas, it is to be expected that the accuracy of the variational solution becomes more critical with respect to the types of trial functions which are used in the calculation.

To show qualitatively the difference between the two trial functions used in the zeroth-order solution and the first-order solution, the constant K appeared in Eq. (33) is compared with the given function $\cos kl$ of Eq. (24) in Fig. 7. Except for certain regions in the neighborhood of the roots of the transcendental $\tan kl = kl$, the constant K determined by a variational procedure behaves almost like the cosine function. One interpretation of this result is that in general the "boundary condition" as defined by Eq. (14) is compatible with the original integro-differential equation given in Eq. (3). There is no significant change in the result if the condition is incorporated in the trial function at the very beginning of the formulation as in the case of designing the zeroth-order solution or if the condition is relaxed as in the case of computing using the first-order solution. In certain critical ranges of kl the result shows that when the "boundary condition" is applied to the trial function at the very beginning of the analysis, we restrict ourselves to a trial function which is not appropriate for these regions. Better results can be obtained only if the "boundary condition" is relaxed so that the function will adjust itself according to the variational procedure to yield a better solution for the problem.

In regard to experimental verification of the theoretical results, there is so far only a limited amount of data on this subject. The measurement reported by D. D. King[8] was based upon the standing-wave method. It discussed the case of broadside incidence with kl varied from 0 to 1.2π, which is about the coverage of the first peak appearing in Fig. 6. As far as this limited data is concerned the variational solution, which is practically the same for various types of trial functions, gives the best agreement with the experimental result among the three methods discussed previously. More experiments,‡ however, should be made in order to check the variational solution as applied to long wires and to verify the angular response curve in case of oblique incidence.

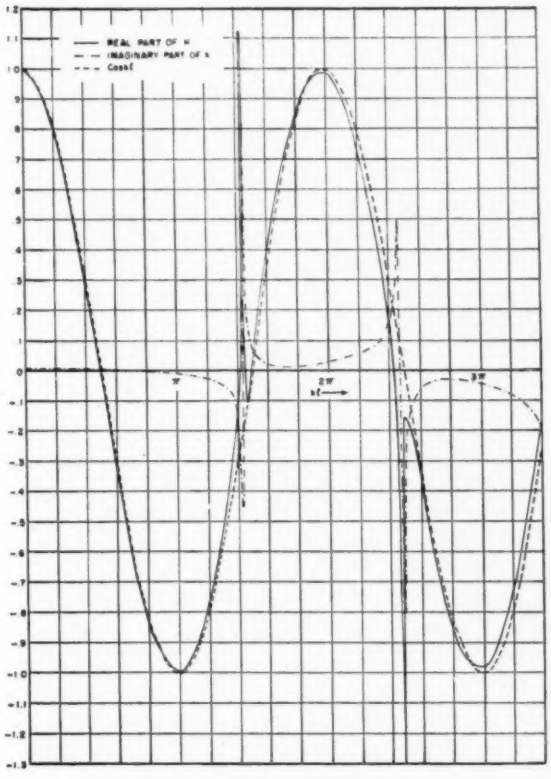


FIG. 7. The parameter K as a function of kl.

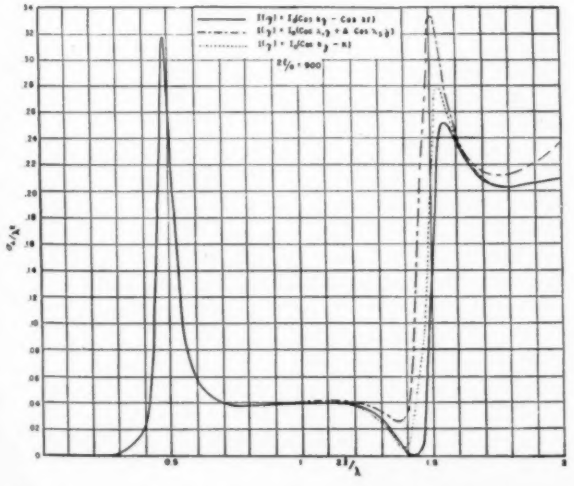


FIG. 6. A comparison of various variational solutions using different trial functions.

ACKNOWLEDGMENT

The author gratefully acknowledges the assistance of Mrs. Elizabeth A. Clarke who carried out the computations of this work.

[8] D. D. King, "Measurement and Interpretation of Antenna Scattering," Cruft Lab. Tech. Rept. No. 50 (Harvard University, Cambridge, Massachusetts, 1948).

‡ *Note added in proof:*—Since the completion of this work, Dr. S. H. Dike and Dr. D. D. King of the Johns Hopkins University have written a number of interesting reports dealing with the cylindrical dipole receiving antenna. The work is contained in Technical Reports Nos. 12, 13, and 14 issued by the Radiation Laboratory, The Johns Hopkins University, Baltimore, Maryland (1951). The subject of back-scattering was extensively discussed in Report No. 12, where the integral equation method was critically examined.

APPENDIX

Formulas Derived from the Variational Method

$$\sigma(\theta)/\lambda^2=3|S|^2/8\pi.$$

$$\frac{1}{S}=\frac{j\int_{-l}^{l}\int_{-l}^{l}I(z)I(z')K(z-z')dzdz'}{\left[k\sin\theta\int_{-l}^{l}I(z)e^{jkz\cos\theta}dz\right]^2},$$

with

$$K(z-z')=k\left(1+\frac{\partial^2}{k^2\partial z^2}\right)\frac{e^{-jkR}}{R};\quad R=[(z-z')^2+a^2]^{\frac{1}{2}}.$$

A. First-Order Solution for Oblique Incidence Using

$$I(z)=I_0[f_c(z)+Af_s(z)].$$

$$\begin{aligned}f_c(z)&=\cos qkl\cos kz-\cos qkz\cos kl,\\ f_s(z)&=\sin qkl\sin kz-\sin qkz\sin kl,\end{aligned}\tag{A1}$$

$$q=\cos\theta,$$

$$g_c=k\sin\theta\int_{-l}^{l}f_c(z)e^{jkz\cos\theta}dz$$

$$=\frac{4q\cos^2qx\sin x-(1+q^2)\sin 2qx\cos x-2q(1-q^2)x\cos x}{2q(1-q^2)^{\frac{1}{2}}},\quad x=kl,\tag{A2}$$

$$g_s=k\sin\theta\int_{-l}^{l}f_s(z)e^{jkz\cos\theta}dz$$

$$=-j\frac{4q\sin^2qx\cos x-(1+q^2)\sin 2qx\sin x+2q(1-q^2)x\cos x}{2q(1-q^2)^{\frac{1}{2}}},\tag{A3}$$

$$\gamma_c=j\int_{-l}^{l}\int_{-l}^{l}f_c(z)f_c(z')K(z-z')dzdz',\tag{A4}$$

$$\gamma_s=j\int_{-l}^{l}\int_{-l}^{l}f_s(z)f_s(z')K(z-z')dzdz'.\tag{A5}$$

When $a\ll l$, γ_c and γ_s are given to a good approximation by the following expressions:

$$\begin{aligned}\gamma_c=&\cos^2x[-1+\cos 2x\cos 2qx+q\sin 2x\sin 2qx\\&-j(\sin 2x\cos 2qx-q\cos 2x\sin 2qx)]\\&-\frac{1}{2}\left(\frac{1+q^2}{q}\cos^2x\cos 2qx+\sin 2x\sin 2qx\right)\\&\times[L(1+q)2x-L(1-q)2x]+j\left\{\cos^2x\right.\\&\left.\times\left[\frac{1+q^2}{2q}\sin 2qx+(1-q^2)x\right]-\sin 2x\cos^2qx\right\}\\&\cdot[\ln 4+\Omega-L(1+q)2x-L(1-q)2x],\end{aligned}\tag{A6}$$

$$\begin{aligned}\gamma_s=&\sin^2x[-1+\cos 2x\cos 2qx+q\sin 2x\sin 2qx\\&-j(\sin 2x\cos 2qx-q\cos 2x\sin 2qx)]\\&+\frac{1}{2}\left(\frac{1+q^2}{q}\sin^2x\cos 2qx-\sin 2x\sin 2qx\right)\\&\times[L(1+q)2x-L(1-q)2x]+j\left\{-\sin^2x\right.\\&\left.\times\left[\frac{1+q^2}{2q}\sin 2qx-(1-q^2)x\right]+\sin 2x\sin^2qx\right\}\\&\cdot[\ln 4+\Omega-L(1+q)2x-L(1-q)2x],\end{aligned}\tag{A7}$$

where

$$x=kl,\quad \Omega=2\ln\frac{2l}{a},$$

$$\begin{aligned}L(1\pm q)2x&=\int_0^{(1\pm q)2x}\frac{1-e^{-ju}}{u}du\\&=\overline{Ci}(1\pm q)2x+jSi(1\pm q)2x.\end{aligned}$$

Expressions (6) and (7) are derived under the condition that when $a\ll l$ the following approximate formulas are valid:

$$\int_0^l\frac{\exp\{-j[(u^2+a^2)^{\frac{1}{2}}-pu]\}}{(u^2+a^2)^{\frac{1}{2}}}du\doteq\ln\frac{2l}{a}-L(1-p)l\tag{A8}$$

$$\int_0^l u\frac{\exp\{-j[(u^2+a^2)^{\frac{1}{2}}-pu]\}}{(u^2+a^2)^{\frac{1}{2}}}du\doteq\frac{-j}{1-p}[1-e^{j(p-1)l}],\tag{A9}$$

where p denotes any given constant.

B. Zeroth-Order Solution for Broadside Incidence Using

$$I(z)=I_0(\cos kz-\cos kl).\tag{A10}$$

$$g_0=g_c|_{q\to 0}=2(\sin x-x\cos x),$$

$$\begin{aligned}\gamma_0=\gamma_c|_{q\to 0}=&2\cos^2x(-1+\cos 2x-j\sin 2x)+L4x\\&+j2\cos x(x\cos x-\sin x)(\ln 4+\Omega-2L2x).\end{aligned}\tag{A11}$$

Thermal Force on an Aerosol Particle in a Temperature Gradient

R. L. SAXTON AND W. E. RANZ
Engineering Experiment Station, University of Illinois, Urbana, Illinois

(Received April 7, 1952)

Data on the thermal forces acting on aerosol particles in a temperature gradient are relatively few. Measurements are reported here on droplets of paraffin oil and castor oil suspended in air, showing the effect of temperature gradient, particle size, and particle material. The thermal force is directly proportional to the temperature gradient prevailing in the gas and to the particle diameter. It is an inverse function of the thermal conductivity of the aerosol material. The agreement between the experimental data and the theory of P. S. Epstein [Z. Physik **54**, 537 (1929)] is satisfactory, a fortunate result in view of the approximations to the actual physical situation made in the derivation of the theoretical equation. The velocity of migration of aerosol particles in a temperature gradient can be estimated from Epstein's formula.

TABLE OF SMYBOLS

D_p = particle diameter, cm
E = voltage gradient, practical volts/cm = $\bar{V}$/plate spacing in cm
F_t = thermal force, dyne
g_c = conversion factor, 980.66 g mass-cm/g force-sec^2
g_L = local acceleration of gravity, 980.14 cm/sec^2
G = temperature gradient prevailing in the gas, deg C/cm
k_a = thermal conductivity of the gas, cal/sec-cm-deg K
k_i = thermal conductivity of the sphere, cal/sec-cm-deg K
K = a constant, equal to 4.5π according to Epstein's theory
L = mean free path of the gas molecules, cm $L=2\mu/\rho\bar{v}$
m = mass of the particle, g mass
p = pressure, dynes/cm^2
q = charge, esu
s = distance along the surface in the direction of increasing temperature, cm
S = length of periphery, cm
T = absolute temperature, deg K
u = gas velocity, cm/sec
v = velocity of the particle, cm/sec
$\bar{v}=(8p/\pi\rho)^{\frac{1}{2}}$, average velocity of gas molecules, cm/sec
$\bar{V}$ = voltage difference, practical volts
W = a group containing all the variables of Epstein's equation except the temperature gradient

$$W=\frac{F_t(2k_a+k_i)\rho T}{D_p k_a \mu^2}, \text{ deg C/cm}$$

$W=KG$
Z = mobility of particle, cm/sec-dyne
$=(1/3\pi\mu D_p)[1+(2L/D_p)\times\{1.23+0.41\exp(-0.44D_p/L)\}]$
ρ = air density, g mass/cc
μ = air viscosity, g mass/cm-sec

Subscripts

a = outside the sphere, or referring to the gaseous medium
i = inside the sphere, or referring to the particle
r = rising
t = thermal in nature

INTRODUCTION

A RADIOMETER force is defined as "a force, other than that caused by convection, which acts on a body suspended in a gas which is not in thermal equilibrium" (34).* It is a force of this type that causes the rotation of the vanes of a radiometer or the migration of small particles suspended in a gas down a temperature gradient.

The earlier observers of the behavior of aerosols in a region of nonuniform temperature were chiefly acquainted with the phenomenon of the "dark space" (1) (32) (38) (39) surrounding a hot body put into an illuminated chamber of dust-filled air. In a series of thoughtfully contrived experiments Aitkin (1) was able to eliminate gravitation, evaporation or combustion, centrifugal and electrical forces, and radiation pressure as possible explanations for the removal of the particles close to a warmed surface. His conclusion was that there exists a force, "thermal repulsion," unique to regions of unequal temperature. Other investigators of the dark space have found that the repulsion does not stem from radiation pressure (7), but that it depends on the pressure of the gas (39) and on the temperature difference between the surface and the gas (39) (29).

More recently, Rosenblatt (34) investigated the effect of pressure on the radiometer force. The velocity of thermal repulsion of single droplets of tricresyl phosphate was measured at a number of temperature gradients and pressures. Particle radii ranged from 0.4 to 1.6 microns, gradients from 8 to 20°C per cm (with a few as high as 50°C per cm), and pressures from 760 to 45 mm Hg. The velocity was independent of particle size, directly proportional to the gradient at constant pressure, and approximately inversely proportional to the pressure when size and temperature gradient were held constant. No significant effect of absolute temperature was observed over the range from 25 to 35°C.

* Numbers in parentheses refer to bibliography at end of article.

In this project it was desired to obtain measurements of the thermal force acting on aerosol particles in a temperature gradient and, more specifically, to determine the several effects of (a) particle size, (b) temperature gradient, and (c) aerosol material on this force. Another objective was to find a relation that would permit thermal forces to be calculated from readily measureable or estimable quantities. This might be found among the existing theoretical developments, in a modification of one of these, or in an empirical correlation.

THEORY OF THE THERMAL FORCE

The mechanism of the thermal or radiometer force is different at high and low pressures, and that at high pressures is the more complex. Here, high pressure means only that the mean free path of the gas molecules is small relative to the dimensions of the body upon which the force acts, while low pressure means that these are of the same order of magnitude. The work of Knudsen (20) (21) (22) (23) (24) elucidating low pressure phenomena preceded the development of the theory of the high pressure region by about a decade. For the present problem of the force on an aerosol particle in a temperature gradient, Knudsen's formulas are of interest only when the particle radius becomes less than twice the mean free path of the gas molecules.

The idea that the radiometer force is caused by a stronger molecular bombardment from the direction of the heat source found expression in Cawood's equation (7):

$$F_t = -\frac{D_p^2}{2}\frac{pL}{T}G. \tag{1}$$

The first approximately correct equation for thermal forces was derived for vane radiometers by A. Einstein (10):

$$F_t = -\frac{S}{2}\frac{pL^2}{T}G. \tag{2}$$

Since L is inversely proportional to p, the equation predicts the experimentally observed pressure dependence of the thermal force, as well as its relation to the temperature gradient (34). It also predicts that the thermal force is directly proportional to a linear dimension of the object on which the force acts and not on an area.

More thoroughgoing derivations have been based on Maxwell's calculation (27) of the stresses set up in an unequally heated mass of gas. At any gas-solid interface molecules are continuously adsorbing on and desorbing from the solid surface. Maxwell deduced that, if a temperature gradient existed along the surface, the desorbing molecules would have a greater component of velocity in the direction of increasing surface temperature than when they arrived at the surface. Thus, the consequence of the interaction of the gas layers close to the surface with the unequally heated solid would be a creeping flow of gas over the surface from colder toward warmer regions. Because this movement arises from a gas-surface interaction, the surface must itself experience a tangential force directed toward the colder regions.

The movement of gas against a temperature gradient along a solid boundary has been named "thermal creep." Its magnitude can be computed from the formula of Maxwell (24)

$$u = -\frac{3}{4}\frac{\mu}{\rho T}\frac{\partial T}{\partial s}. \tag{3}$$

Reynolds' (33) expression for the thermal stress was identical in form with Maxwell's but only one-third as great. Hettner (16) proceeded from Maxwell's principles to the concept that radiometer forces in the region of normal gas densities are produced by thermal creep. The temperature differences established on the surface of the system by external effects, e.g., by heat conducted through the gas or by intercepted radiation, lead to a slippage of the adjoining gas layers in the direction of increasing temperature. The reaction of the body to this movement is observed as the thermal force. That the gas currents traced by Hettner actually exist has been proved experimentally by Gerlach and Schutz (4).

The most comprehensive development of radiometer theory is due to Epstein (12). This investigator not only recognized the kinetic and hydrodynamic aspects of the

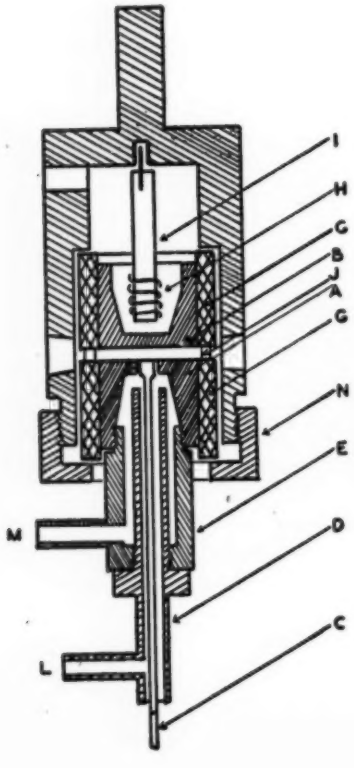


FIG. 1. Diagram of observation cell. *A*—lower plate, *B*—upper plate, *C*—centering device, *D*—water inlet tube, *E*—water outlet tube, *G*—insulating sheath, *H*—heater, *I*—alundum cylinder, *J*—plate spacers, *L*—water inlet, *M*—water outlet.

problem, but also took account of the fact that the behavior of a body in a gas depends on the ratio of heat transported through its interior (internal conduction) to the heat received in unit time from molecular impacts (external conduction) (11). Therefore, the thermal conductivities of the gas and the immersed material should enter into the derivation of the thermal stress.

In his derivation for the case of a sphere Epstein assumed (a) that the mean free path was small relative to the dimensions of the object experiencing the force, (b) that a uniform temperature gradient existed in the gas at a great distance from this object, and (c) that Fourier's equation for heat conduction without convection could be used even though such a state could not exist because of the thermal creep. He then broke down the problem into the two phases of hydrodynamics and heat conduction which were treated independently at first, but later combined through the concept of the thermal creep.

Epstein's procedure may be summarized as follows. The heat conduction problem was solved to obtain the temperature gradient over the surface of the sphere in terms of that prevailing in the gas. From this could then be calculated the velocity of the thermal creep. Next the Navier-Stokes equations of motion were solved with the assumption that the inertia terms are negligible and with the creep velocity at the surface of the sphere as a boundary condition. The thermal force was then found as the integral over the surface of the component of stress parallel to the direction of heat flow.†

The final expression for the thermal force, as derived by Epstein, is

$$F_t = -9\pi(D_p/2)(\mu^2/\rho T)[k_a/2k_a+k_i]G, \qquad (4)$$

where $3[k_a/2k_a+k_i]G$ represents the calculated temperature gradient in the vicinity of the particle.

TEST EQUIPMENT AND EXPERIMENTAL PROCEDURE

1. Equipment

The method used in measuring the thermal force was an adaptation of Millikan's familiar oil-drop experiment. For each particle the times of free fall and of rise under a known electrostatic field were recorded so that the mass and charge could be computed. Then a temperature gradient was established in the air between the plates, the upper plate being made the hotter so as to insure the absence of convection. The strength of the electrostatic field needed to lift the particle against the temperature gradient was determined, and the thermal force was obtained as the difference between this electrical force and the sum of gravity and drag forces:

$$F_t = (qE/300) - mg_L - (v_r/Z). \qquad (5)$$

† For a review of this derivation see R. L. Saxton and W. E. Ranz, Technical Report No. 6, SO-1007, Contract No. AT(30-3)-28, United States Atomic Energy Commission.

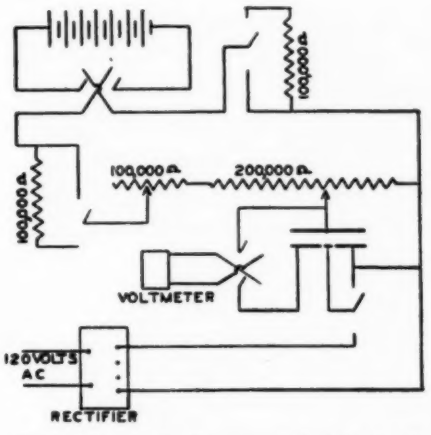


FIG. 2. Schematic diagram of electric circuit.

The experimental conditions were restricted to air at atmospheric pressure and approximately room temperature.

The characteristics of an aerosol material suitable for this work were (a) that it be stable and nonvolatile at temperatures up to 60°C, (b) that it be of known thermal conductivity, and (c) that it be easily atomized so that spherical particles could be obtained. The two substances chosen were castor oil and a high molecular weight paraffin oil which had thermal conductivities of 4.32×10^{-4} and 2.97×10^{-4} cal/cm-sec-deg C, respectively at 20°C (36). The corresponding figures at 40°C were 4.28×10^{-4} and 2.94×10^{-4} cal/cm-sec-deg C; values for intermediate temperatures were found by linear interpolation. The ratio of these conductivities, 1.45, was thought sufficient to demonstrate the effect of thermal conductivity.

An aerosol was formed by atomizing oil within a glass bulb by a jet of dried air. The aerosol particles were then electrically charged by being passed through a corona discharge.

After receiving an electrical charge the droplets entered the cell shown in Fig. 1 where a number of them were captured and held suspended between condenser plates *A* and *B* which were approximately 0.3 cm apart. In the space above the top plate was fixed a coil of chromel resistance wire *H*. The current to this small heating unit was regulated by a Variac and a rheostat connected in series so that the temperature of the top plate could be controlled. The inner lip of the lower plate *A* was threaded for coupling to the water supply assembly *D*. Cooling water entered at *L* and flowed up between the tube *D* and the rod *C* to bathe the underside of the plate *A*. The brass rod *C* ran up through the tube *D* and flared into a plug. This plug fitted into a hole at the center of the lower plate *A*, and its surface lay flush with the plate face. Rod *C* was insulated from *A* and from the cooling water. Such an arrangement was devised by Fletcher (12) to permit recentering particles whenever they drift out of the field of view.

The condenser plates were charged by a bank of 67.5-

volt "*B*" batteries. From these batteries the voltage connections passed through a reversing switch and two carbon potentiometers in series as shown in Fig. 2. This arrangement gave a vernier-like control over the voltage tapped off for charging the plates. The voltage difference across the plates was registered continuously on a voltmeter. The voltages required to suspend the particles in the presence of a temperature gradient ranged up to 300 volts, corresponding to a voltage gradient of 915 volts per centimeter in the air gap.

A microscope light powered by a variable transformer supplied light for the cell. To remove the longer wavelengths, which produce the greatest photophoretic and heating effects, this light was filtered through a foot of dilute copper sulfate solution. It then entered the cell at an angle of 150 degrees to the line of view.

The particles were observed through a microscope mounted on a traversing device. Under twentyfold magnification they appeared as bright, star-like points against a relatively dark background. The microscope eyepiece contained a reticule which was calibrated against a stage micrometer; two divisions at the center of the disk were equal to one millimeter at the focus of the microscope. The velocities of particles were measured by timing their movement over this distance with a stopwatch.

Two chromel-alumel thermocouples used to measure plate temperatures were calibrated against Bureau of Standards thermometers that were correct to 0.01 degree. Temperatures were read from the calibration curves, and were considered accurate to 0.1°C. When in place, each thermocouple fitted into a well, drilled parallel to and $\frac{1}{16}$ inch below the plate surface. The thermocouples were assumed to read surface temperatures, the drop across the thin layer of metal separating them from the plate surface being neglected. Since the temperature gradient in the space between two close spaced, parallel plates has been shown to be linear (31); the temperature gradient was then obtained by dividing the temperature difference by the plate separation.

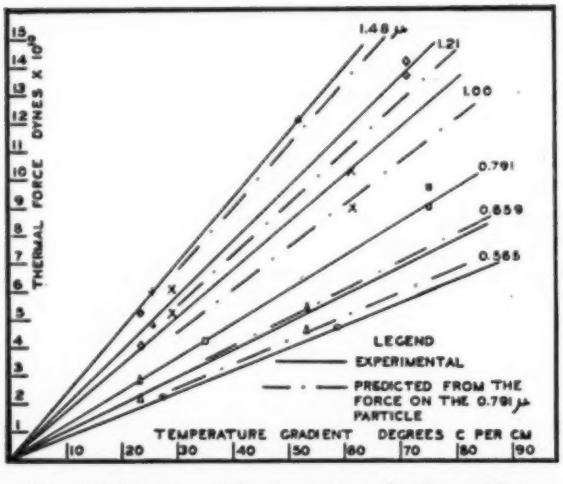


FIG. 3. Effect of particle diameter on the thermal force.

2. Procedure

In a typical experiment, the atomizer, charger, and cell were first purged with a slow stream of dry air. After a few minutes the voltage to the charger was set at the working value, the air velocity in the atomizer was increased, and vacuum was applied to the cell outlet. Air and vacuum were adjusted until a slow stream of finely dispersed oil particles was being generated. After about a minute air and vacuum were turned off, and the cell petcocks closed. Next, the battery circuit was closed and the voltage across the plates increased slowly until a particle or particles of the desired size were suspended.

Particle size could be estimated readily after some experience, and one particle could be chosen for study. A precaution that was observed in selecting particles for long observation was that the charge to mass ratio was not so large that small fluctuations in voltage would cause the particle to be precipitated. Another disadvantage inherent in a disproportionately high charge was that only a small change in field strength was required to compensate for the presence of a temperature gradient; consequently, in the calculation of the thermal force, a large constant factor, the charge, would be multiplied by a variable, the field strength, which changed only a little. This, of course, would introduce a greater uncertainty in the calculated force.

By running the particles up and down in the cell and applying the centering device, the cell was cleared of all but the desired particle. Usually ten falling times were then measured, and five times of rise under each of two different fields. The average values were used to compute the falling and rising velocities, from which in turn the mass and charge of the droplet were calculated.

After these data had been obtained, a temperature gradient was established in the air gap. The sudden temperature change caused the particle to jump around violently; but after the plate temperatures had been constant for about 20 minutes, reproducible measurements of rising or falling velocities could be obtained. At temperature gradients above 40°C per cm the time of rise under a given field strength was observed to diminish steadily upon repeated measurement, as if the thermal force were decreasing. This was ascribed to the mixing action of the particle itself. As the particle rose and fell it would drag warm air down and cold air up, and since it traveled in very nearly the same path on each trip there would be created a little cylinder within which the temperature gradient was partially destroyed. When a few minutes were allowed to elapse between trips, check readings could be obtained. Again the average of a number of times was used to compute the velocity.

Usually two, and occasionally three, temperature gradients were studied for each particle. The particle was held under virtually continuous observation during this time, from five to eight hours in duration.

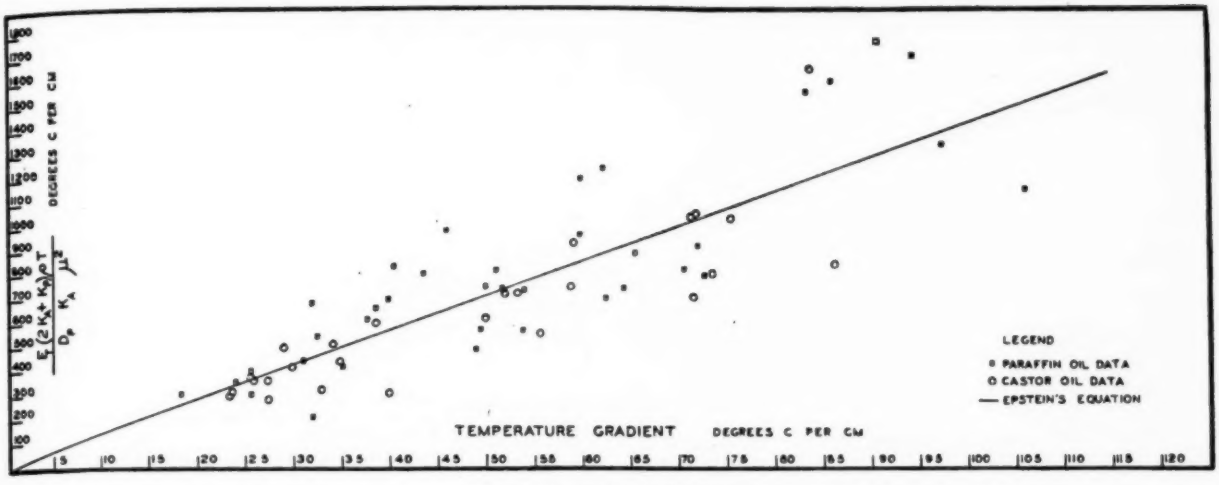


FIG. 4. General correlation.

DISCUSSION OF RESULTS

Of the data for 55 particles studied, the results for 25 paraffin and 15 castor oil particles were retained. The basis for rejection, outside of an unsatisfactory experiment, was a marked deviation from linearity of the plot of thermal force *versus* temperature gradient. Particle diameter ranged from 0.481 to 2.16 microns, and temperature gradients from 18.4 to 140°C per cm. The thermal forces corresponding to these gradients were 2.07×10^{-10} dyne ($D_p=0.481$ micron), and 65.9×10^{-10} dyne ($D_p=1.36$ microns), respectively. These may be compared with the photophoretic forces on sulfur and selenium particles measured by Parankiewicz (30) and Mattauch (26) which were of the order of 10^{-10} to 10^{-11} dyne. The average atmospheric pressure was 745 mm Hg, and the temperature in the observation space, taken as the average of the plate temperatures, was about 27°C.

1. Effect of Particle Size

The thermal force was found to be proportional to particle diameter, as stated by Epstein's equation. This is illustrated in Fig. 3 in which the experimental forces (solid lines) for six castor oil particles are compared with those predicted from a proportionality of force and diameter (dashed lines). The 0.791-micron particle serves as the basis for the calculation; the dashed lines are placed according to the product of the measured force on this particle and the ratio $D_p/0.791$. The exceptions to this rule generally fell within the experimental error.

2. Effect of Temperature Gradient

Figure 3 also shows that for a given particle size, the thermal force is directly proportional to the temperature gradient, again in agreement with theory, and with previous experiments (34). As pointed out by Rosenblatt (34), a change in cell temperature of about seven degrees could be tolerated before the effect of absolute temperature became significant.

3. General Correlation

Figure 4 represents a general correlation that takes the form

$$W=KG, \tag{6}$$

where $W=F_t(2k_a+k_i)\rho T/D_p k_a\mu^2$, degrees C per cm, and K is a constant.

The graph demonstrates that the combination of variables proposed by Epstein serves to correlate the experimental data; it remains to evaluate the constant K. Since all variables are combined in the group W, each point on the graph now represents an independent measurement. Hence, K can be evaluated for each point, and these values compared with Epstein's figure, $4.5\pi=14.14$.

For the paraffin oil particles the average individual deviation of K from the theoretical value was 21.2 percent. The average value of K from this series of measurements was 14.73, differing from 4.5π by 4.17 percent. For the castor oil particles the average individual deviation was 14.5 percent, and the average K was 14.29, only 1.06 percent above the theoretical constant. The over-all average value of K, determined from the measurements of both series, was 14.1 which is only

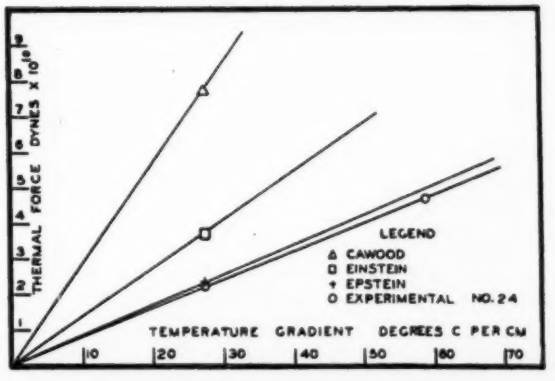


FIG. 5. Comparison of theoretical equations.

0.28 percent removed from Epstein's value. It would appear that the deviations of individual experiments were random so that, on the average, the experimental data confirm the theoretical relation.

4. Comparison with Other Theories

The apparent superiority of Epstein's equation over those of Cawood and Einstein is illustrated by Fig. 5 in which the calculated thermal force is compared with a measurement considered accurate. The force predicted by the equations of Cawood, Einstein, and Epstein differed from the experimental value by 249, 68.5, and 5.4 percent, respectively.

5. Experimental Error

Besides the usual inaccuracies involved in reading instruments, timing the movement of particles, etc., errors could arise from convection in the observation space, evaporation of the droplets, or fluctuation in the charge on the particle. Those errors which could be estimated (including those for voltage, temperature, pressure, and time) were subjected to error analysis. For a 1.47-micron paraffin oil droplet, under a temperature gradient of 34.2°C per cm, these calculations indicated a maximum error in the thermal force of 29.5 percent and a minimum error of 9.15 percent. For the same particle the estimated maximum and minimum error at a gradient of 106°C per cm were 16.7 and 4.66 percent, respectively. A reasonable estimate of the probable error in the measured thermal forces was 20 percent, being approximately equal to the average individual deviation of K from the theoretical value.

Evaporation was proved to be negligible by mass rechecks. The diameter of a 0.604-micron paraffin oil particle held under observation for 6½ hours remained constant to 0.5 percent, and the diameter of a 0.657-micron castor oil droplet that was suspended in the cell for seven hours checked the original measurement to 3.5 percent.

The charges of the two particles for which the mass was checked were redetermined and found unchanged, but the charges on several other droplets were observed to change while they were being measured. A change in charge that occurred after a temperature gradient was set up in the cell would almost certainly have gone unnoticed. However, for most of the particles studied, including the two for which the charge was remeasured, the thermal force was linear in the temperature gradient. Hence, if a marked nonlinearity was apparent in the curve of F_t *versus* G, the data were not used in the overall correlation.

The absence of convection was affirmed by the stability of the suspended particles once a temperature gradient had been established. Paranjpe (25) could detect no convection in a similar observation space at plate spacings below 0.5 cm.

Another uncertainty arises in the actual magnitude of the mobility of the particles. This was calculated from Stokes' law (37) with a correction factor added by Millikan (28). The correction factor is analogous to that derived by Cunningham (9) from kinetic considerations. Its purpose is to take account of the discontinuous nature of the medium for particles having diameters of the order of magnitude of the mean free path. For such particles the medium is said to "slip" over the surface, and the velocity at the surface is no longer zero, as assumed in Stokes' derivation. However, it is implicitly assumed in both Cunningham's and Millikan's corrections that the sole cause of slip in gases is found in inhomogeneities on a molecular scale (28). If this were not true, Stokes' unmodified law would not hold accurately in dense media, as it is experimentally found to do.

To this picture must now be added the concept of the thermal creep, also a slippage of gas at the surface of an immersed body, but one that is probably independent of the movement of the particle for a constant temperature gradient. The effect of this slippage on the resistance to movement of the particle cannot readily be ascertained.

Epstein's relation for the thermal force, and the corrected form of Stokes' law both stem from solutions to the same hydrodynamic problem, but the boundary conditions assigned are contradictory. In determining the thermal force experimentally the particle mobility calculated from the modified Stokes' law was used. The resulting value of the thermal force was then compared with that predicted by the theory of Epstein, and these were found to agree within the limits of experimental error. It can only be concluded that, despite this evident inconsistency, the velocity of migration of aerosol particles under the influence of a temperature gradient can be calculated from the thermal force given by Epstein's relation, and the particle mobility can be computed from a corrected version of the law of Stokes, the two quantities being apparently independent of one another.

CONCLUSIONS

The result of this investigation has been to substantiate Epstein's equation for radiometer forces. This relation, Eq. (4), states that the thermal force on an aerosol particle in a temperature gradient is directly proportional to this gradient, and to the particle diameter. In addition, the force is shown to be an inverse function of the thermal conductivity of the aerosol material. The agreement between theory and experiment seems satisfactory, the more so in view of the approximations to the actual physical picture that were made in the derivation.

Finally, it appears that the velocity of migration of aerosol particles down a temperature gradient can be computed from the thermal force as given by Epstein's equation, and a particle mobility based on a corrected form of Stokes' law.

ACKNOWLEDGMENT

This work has been carried out in the Engineering Experiment Station of the University of Illinois as part of a contract for aerosol studies, Contract No. AT(30-3)-28, United States Atomic Energy Commission, H. F. Johnstone, Technical Director. Detailed information concerning the theory and compilations of the experimental data are available as Technical Report No. 6, SO-1007, Technical Information Service, United States Atomic Energy Commission, Oak Ridge, Tennessee.

BIBLIOGRAPHY

(1) J. Aitken, "On the formation of small clear spaces in dusty air," Trans. Roy. Soc. Ed. **32**, 239 (1884).
(2) W. D. Bancroft, "Thermal filters," J. Phys. Chem. **24**, 421 (1920).
(3) S. C. Blacktin, "The cleaning of air and gas by thermal repulsion," J. Soc. Chem. Ind. (London) **58**, part 2, 334 (1939).
(4) S. C. Blacktin, J. Soc. Chem. Ind. (London) **59**, part 3, 153 (1940).
(5) J. Bredl and T. W. Grieve, "A thermal precipitator for the gravimetric estimation of solid particles in flue gases," J. Sci. Inst. **28**, 21 (1951).
(6) C. F. Casella, Ltd., Fitzroy Square, London, W 1, Leaflet 720 *A*.
(7) W. Cawood, "The movement of dust or smoke particles in a temperature gradient," Trans. Faraday Soc. **32**, 1068 (1936).
(8) W. Crookes, "On repulsion from radiation," Trans. Roy. Soc. (London) **170–1**, 87 (1879).
(9) E. Cunningham, "On the velocity of steady fall of spherical particles through fluid medium," Proc. Roy. Soc. (London) **83A**, 357 (1910).
(10) A. Einstein, "Zur theorie der radiometerkrafte," Z. Physik **27**, 1 (1924).
(11) P. Epstein, "On resistance experienced by spheres in their motion through gases," Phys. Rev. **23**, 710 (1924).
(12) P. Epstein, "Zur theorie des radiometers," Z. Physik **54**, 537 (1929).
(13) H. Fletcher, "A determination of Avogadro's constant *N* from measurements of the Brownian movements of small oil drops suspended in air," Phys. Rev. **39**, 446 (1914).
(14) W. Gerlach and W. Schutz, "Untersuchungen an radiometern IV," Z. Physik **78**, 43 (1932).
(15) H. L. Green and H. H. Watson, "Physical methods for the estimation of the dust hazard in industry," Medical Research Council, Special Report Series, No. 199, London H.M.S.O. (1935).
(16) G. Hettner, "Zur theorie des radiometers," Z. Physik **27**, 12 (1924).
(17) G. Hettner, "Zur theorie der photophorese," Z. Physik **37**, 179 (1926).
(18) J. C. Hunsaker and B. G. Rightmire, *Engineering Applications of Fluid Mechanics* (McGraw-Hill Book Company, Inc., New York and London, 1947), p. 144.
(19) J. E. Jones, "On the velocity distribution function and on the stresses in a non-uniform rarefied monatomic gas," Phil. Trans. **223**, Series *A*, (1923).
(20) M. Knudsen, "Eine revision der gleichgewichtsbedingung der gase thermische molekularstromung," Ann. Physik **31**, 205 (1910).
(21) M. Knudsen, "Thermischer molekulardruck der gase in rohren und porosen korpern," Ann. Physik **31**, 633 (1910).
(22) M. Knudsen, "Ein absolute manometer," Ann. Physik **32**, 809 (1910).
(23) M. Knudsen, Ann. Physik **33**, 1435 (1910).
(24) M. Knudsen, "Erwiderung an Hrn. M.v. Smoluchowski," Ann. Physik **34**, 823 (1911).
(25) H. Lamb, *Hydrodynamics* (Dover Publications, New York, 1945), sixth edition, p. 597.
(26) J. Mattauch, "Neue versuche zur photophorese," Physik Z. **23**, 444 (1922).
(27) J. C. Maxwell, "On stresses in rarefied gases arising from inequalities of temperature," Trans. Roy. Soc. (London) **170**, 231 (1879).
(28) R. A. Millikan, "Coefficients of slip in gases and the law of reflection from the surfaces of solids and liquids," Phys. Rev. **21**, 217 (1923).
(29) M. K. Paranjpe, "The convection and variation of temperature near a hot surface. The dust-free dark layer in relation to surface convection," Proc. Indian Acad. Sci. **4a**, 423 (1936).
(30) I. Parankiewicz, "Uber die lichtpositive und die lichtnegative photophorese (untersucht am Schwefel und Selen)," Ann. Physik **57**, 489 (1918).
(31) L. A. Ramdas and M. K. Paranjpe, "An interferometric method of measuring temperatures and temperature gradients very close to a hot surface," Current Sci. (India) **4**, 642 (1936).
(32) Lord Rayleigh, "On the dark plane which is formed over a heated wire in dusty air," Proc. Roy. Soc. (London) **34**, 414 (1882).
(33) O. Reynolds, "On forces caused by communication of heat between a surface and a gas," Trans. Roy. Soc. (London) **166**, 725 (1876).
(34) P. Rosenblatt and V. K. Lamer, "Motion of a particle in a temperature gradient; thermal repulsion as a radiometer phenomenon," Phys. Rev. **70**, 385 (1946).
(35) T. Sexl, "Zur theorie des einplattenradiometers," Z. Physik **52**, 249 (1928).
(36) J. F. D. Smith, "The thermal conductivity of liquids," Trans. Am. Soc. Mech. Engrs. **58**, 719 (1936).
(37) G. G. Stokes, "On the effect of the internal friction of fluids on the motion of pendulums," Trans. Cambridge Phil. Soc. **9** (part 11), 8 (1850).
(38) J. Tyndall, "On dust and disease," Proc. Roy. Inst. **6**, 3 (1870).
(39) H. H. Watson, "The dust-free space surrounding hot bodies," Trans. Faraday Soc. **32**, 1073 (1936).

Free Molecular Flow in the Sample Inlet to the Mass Spectrometer

PAUL D. ZEMANY
General Electric Research Laboratory, Schenectady, New York
(Received April 10, 1952)

The aspects of free molecular flow in the sample inlet to the mass spectrometer are pointed out and the conditions under which such flow can occur are described. Methods for calculating the leak diameter and sample pressure for various flow rates are developed. These principles are illustrated by consideration of the flow characteristics in the mass spectrometer leak detector, and by an example from reaction kinetic studies.

The decrease in pressure in the sample reservoir is also described. This is applied to a determination of molecular weights. The effect of the pressure decrease on apparent fragmentation patterns is pointed out and a rational expression for spectrometer sensitivity is suggested which is based on the fraction of molecules which flow through the spectrometer which appear as ions.

INTRODUCTION

IT was shown by Honig[1] that there is free molecular flow[2] in the mass spectrometer beyond the inlet leak (Fig. 1), but he placed less emphasis on the conditions necessary in the sample reservoir and in the leak itself. Since the flow rate through the spectrometer is completely determined by this part of the system, it will be important to scrutinize these flow conditions carefully for all possible cases.

The flow rate through this leak must be a well-defined function of the pressure of gas in the sample system if quantitative results are to be obtained on the spectrometer. This well-defined functionality exists in either of two limiting cases: First, when the flow rate is proportional to the partial pressure of each gas in the sample reservoir (molecular flow); second, when this rate is proportional to the square of the total pressure of the gas mixture[3] (viscous flow). The conditions for molecular flow, the simplest case, will be elucidated below and a number of applications possible in this case will be discussed. Viscous flow will be mentioned only in connection with the leak detector and the complex region intermediate between these two limiting cases will not be discussed at all. There is a third variety of gas flow, called turbulent flow, which is not likely to be encountered in applications of the mass spectrometer and which will not be mentioned further.

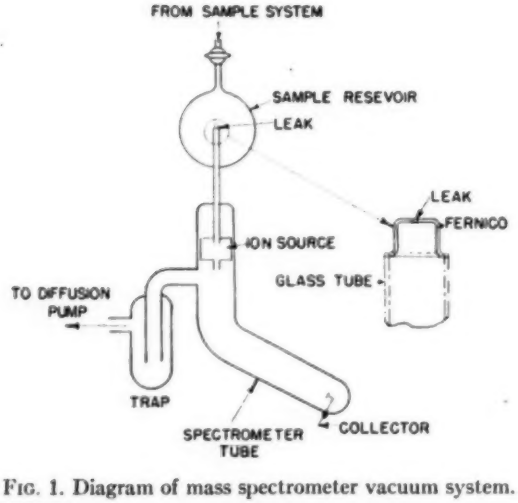


FIG. 1. Diagram of mass spectrometer vacuum system.

Molecular flow is the process of the effusion of a gas through an orifice. While some of the mathematical expressions for this process are similar to those which describe Graham's law, the two processes should not be confused. Graham's law applies to the diffusion of gases.

The rate of flow through the leak, Fig. 1, determines the peak height observed on the mass spectrometer and there is a direct proportionality between the two. This rate of flow is also directly proportional to the partial pressure of the gas in the sample volume only under conditions where molecular flow is occurring. Knudsen showed that a practical criterion which is now commonly accepted[4] for free molecular flow be that the diameter of the orifice d be at most $\frac{1}{10}$ the mean free path L of the gas at the higher pressure. With the pressure such that the mean free path was less than ten times the orifice diameter the effusion rate increased, and with a further increase in pressure the rate gradually shifted over to viscous flow. At $L=2d$, the error amounted to a few percent for some gases. The factor of ten is an empirical one that seems to offer an adequate margin of safety. For rough mass spectrometer quantitative work, this may be too large; for qualitative work, the restriction is unnecessary.

The rate of gas flow into the spectrometer, which is determined solely by the leak, has to fulfill certain other conditions. It is fixed on the low side by the total sensitivity of the spectrometer, and on the high side by the maximum operating pressure. Other special limiting factors might also be present. It will be of some interest, then, to determine the relations between rates of gas flow, orifice diameter, and pressure that limit the range of molecular flow.*

[1] R. E. Honig, J. Appl. Phys. **16**, 646 (1945).
[2] M. Knudsen, Ann. Physik **25**, 983 (1908); **28**, 75, 999 (1909).
[3] R. E. Halstead and A. O. Nier, Rev. Sci. Instr. **21**, 1019 (1950).
[4] S. Dushman, *Scientific Foundations of Vacuum Technique* (John Wiley and Sons, Inc., New York, 1949), p. 93.

* Most of the following discussion is based on the treatment given in reference 4.

REQUIREMENTS FOR MOLECULAR FLOW THROUGH THE LEAK

The volume F of gas which strikes unit area in unit time is $F=V_a/4=3638(T/M)^{\frac{1}{2}}$cm^3 sec^{-1} cm^{-2} where V_a is the average velocity, T the absolute temperature, and M the molecular weight. This is the volume that would effuse through an orifice, a square centimeter in area, in one second. For a circular orifice of diameter, d cm, this becomes $F=2.86d^2(T/M)^{\frac{1}{2}}$ liters sec^{-1} on substituting the proper constants.

The quantity of gas that goes through the leak each second will be $Q=2.86d^2p(T/M)^{\frac{1}{2}}$ micron liters where p is the pressure of gas in the gas reservoir expressed in microns. (It is assumed that the pressure in the spectrometer itself is so low that back diffusion is negligible.) The maximum desirable rate of gas flow into the spectrometer depends upon many factors, which may differ depending upon the application. In the spectrometer used, a rate of about 0.01μ liter per second of CO_2 is well within the working pressure of the spectrometer, but gives several times full-scale deflection of the recorder on the least sensitive scale. This rate will be arbitrarily adapted for purposes of illustration.

The mean free path of a gas, L_0, depends inversely on the pressure and the square of the molecular diameter δ, and is given by $1/L_0=\sqrt{2}\pi n\delta^2$ where n is the number of molecules per cm^3. There is no simple expression for molecular diameter, but the values for various gases do not vary over a very wide range: (He=2.2A, CO_2=4.6A, benzene=8.2A). Carbon dioxide may again be used as an example. At one micron pressure and 25°C, $L_0=3.34$ cm. At any other pressure p, expressed in microns, $L=3.34/p$ cm.

Having specified that the rate of flow of CO_2 be 0.01μ liter sec^{-1}, and that the diameter of the orifice be, at most, $\frac{1}{10}$ the mean free path, it is possible to determine the values of the pressure that would be required for various sizes of orifice to fulfill this condition. With $Q=0.01$ micron liter per second $=2.86d^2p(T/M)^{\frac{1}{2}}=7.40d^2p$ and $L=10d=3.34/p$ cm for CO_2, it is found that $d=0.00405$ cm, $p=82.5\mu$ as the minimum leak diameter and the maximum pressure allowable. The same rate of flow could be achieved by using a larger leak, and correspondingly lower pressure, but if a smaller leak were used, an attempt to get a flow of 0.01 micron/liters per second by increasing the pressure would result in a departure from free molecular flow.

With the same leak diameter of 0.00405 cm, the maximum pressure of helium, which has a mean free path of 14.72 cm/p, is 370μ and the maximum rate of gas effusion will be 0.149μl/sec. Benzene, on the other hand, with a mean free path of $1.13/p$ cm will have a maximum pressure of 27.9 and effusion rate of 0.00255 micron liters/second.

As another example, in some reaction kinetic studies[5] it was desired that a sample of gas be pyrolyzed in a combination reaction vessel-leak reservoir at a pressure in the centimeter range; and that the composition of the gas be measured continuously by mass spectrometer recording. The gas, acetaldehyde, with a molecular weight of 44 may be assumed to have a somewhat larger diameter than CO_2, and a mean free path of, say, 3.0 cm at 1μ and 25°C. At a maximum operating pressure of 1.5 cm or $15{,}000\mu$ the mean free path is 2×10^{-4} cm, the orifice diameter 2×10^{-5} cm; the rate of flow will then be $4.2\times10^{-5}\mu$l/sec. The actual flow rate was about five times this value at 500°C; however, L, the mean free path, increases with temperature, thus permitting an increase in the orifice diameter while preserving molecular flow. Also, the flow rate will increase with temperature. So with the pressures and temperatures used, the conditions for molecular flow were fulfilled.

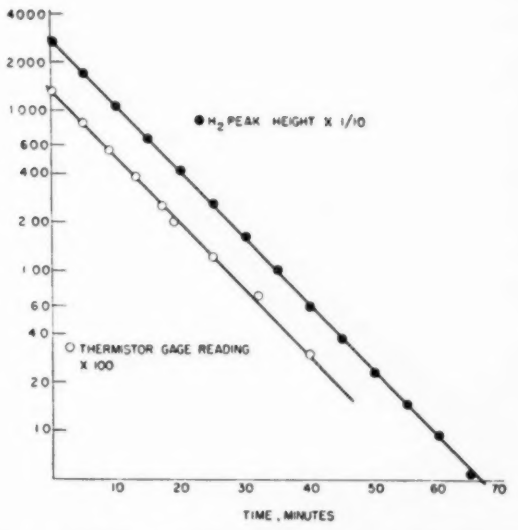


FIG. 2. Effusion of hydrogen from sample reservoir. Peak heights indicate rate at which hydrogen passed through the mass spectrometer ion source; thermistor gauge readings give the pressure of hydrogen remaining in the sample reservoir.

EXPERIMENTAL VERIFICATION OF MOLECULAR FLOW THROUGH A LEAK

The data presented in Fig. 2 show both the pressure of hydrogen and the hydrogen peak height as a function of time. The hydrogen pressure was measured by a thermistor pressure[6] gauge in the one-liter sample reservoir. The hydrogen peak height was measured continuously on the same sample as it effused through the leak. This leak was of the design of F. J. Norton,[7] made by prick punching a small hole in a 1-mil Fernico diaphragm, peening it over, and etching with acid to give the desired leak rate. This leak projected to the center of the sample reservoir, as in Fig. 1, and was connected to the mass spectrometer by about 40 cm of 0.8-cm diameter tubing. A General Electric analytical

[5] P. D. Zemany and M. Burton, J. Chem. Phys. **55**, 949 (1951).

[6] P. D. Zemany, Rev. Sci. Instr. **23**, 176 (1952).

[7] F. J. Norton (private communication).

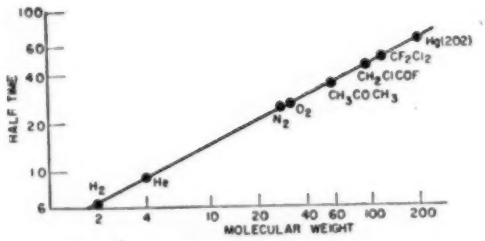


FIG. 3. Log–log plot of time required for peak height to decrease to half its initial value *vs* molecular weight for several gases, showing a linear relationship with a slope of one-half.

mass spectrometer was used. From the figure it may be observed that

$$d \log p_t/dt = d \log p_h/dt = K_1,$$

where p_t is the pressure in microns, measured by the thermistor gauge, p_h the peak height observed in the mass spectrometer, and K_1 one form of the leak constant. Since both slopes are constant over the entire pressure range, the leak rate is constant. Thus, the leak exhibits one property of molecular gas flow. Since the ratio p_t/p_h is constant, the ion current is directly proportional to the gas pressure in the sample reservoir over the entire pressure range, as required.

In Fig. 3, there are plotted the logarithms of the half-times of gases of various molecular weights against the logarithms of the molecular weights. The slope of the line is 0.50, hence the rate of effusion through the leak is inversely proportional to the square root of the molecular weight, another necessary condition for molecular flow.

These experiments were performed at room temperature (25°) and it was not convenient to alter this for measuring the leak rate as a function of temperature; however, other experiments[6] with a similar leak show the rate proportional to the square root of the absolute temperature, within the accuracy of the measurement. Thus, this leak exhibits all the properties of free molecular flow.

GAS FLOW IN MASS SPECTROMETER LEAK DETECTOR

Another important calculation is that of the maximum diameter and flow rate for helium at one atmosphere of pressure, which is of interest in connection with the mass spectrometer leak detector. At 25°C and 760-mm pressure, the mean free path of helium is 19.36×10^{-6} cm; hence, the diameter of the leak must not exceed 1.936×10^{-6} cm. This corresponds to a leak rate of 7.5×10^{-5} micron liter per sec. Any actual leak detected by the mass spectrometer leak detector (the limit of sensitivity of which is in the order of $5\times10^{-7}\mu$l/sec), which exceeds this rate will tend to be by viscous flow. Since most leaks encountered are more likely to be long, thin fissures rather than holes in thin plates, the region of viscous flow extends to much lower leak rates. Thus, most leaks detected on the mass spectrometer leak detector exhibit viscous flow; however, once the probe gas is in the low vacuum side, specifically in the ionization chamber, it will obey the laws of free molecular flow.

To show this, a leak which had a rate of $2.3\times10^{-5}\mu$l/sec of air was attached to the mass spectrometer. Peak heights with helium and argon were compared when each was exposed to the leak at one atmosphere pressure. The instrument used here was also the General Electric analytical mass spectrometer. It was found that the argon response was 21 times that of helium. The ratio of responses, through a leak which exhibits molecular flow, is 10:1. Hence, the argon flowed through the viscous leak 2.1 times as readily as it would had the flow been molecular. One would have expected, on the basis of a limiting viscous flow[8] through the leak and free molecular flow through the ion source, a ratio of 2.8. Since the limit of flow through a leak is governed by the laws of viscous flow, the use of helium as a probe gas, because of its small mass, is of no advantage. Argon (as an inert gas) would give greater sensitivity. It would also have the advantage of a higher ionization efficiency. The ratio 10:1 obtained above is greater than the value 7:1 obtained by Young and Dushman[9] for the relative ionization efficiencies, probably on account of more efficient focusing of the heavier ion.

Increasing the probe gas pressure would also yield a more than proportional increase in sensitivity, since the flow rate through the leak increases with the square of the pressure.

RATE OF DECAY OF PRESSURE IN THE SAMPLE RESERVOIR

Another important consideration is the fact that the pressure of the sample in the reservoir decreases with time. This has several consequences. It limits the time during which a sample can be observed. In the reaction kinetics study referred to, the rates of pressure decrease were in the order of a few hundredths percent an hour. It had been deliberately made as slow as was practical. Ordinarily it amounts to a few tenths percent per minute. It leads to fractionation of the sample as Honig[1] pointed out. It may be a source of discrepancies between fragmentation patterns observed on different instruments as will be shown below. It offers a method of estimating the molecular weight of a substance by using the mass spectrometer as a pressure gauge.[10]

A brief discussion of the mode of pressure decay is given in a previous paper.[6] Figure 2 shows a typical plot of log p *vs* time, and Fig. 3 the half-times for a variety of gases effusing through a leak from a sample reservoir of one liter volume. The half-time for any gas is $4.6M^{\frac{1}{2}}$ minutes with this particular leak and sample volume. The value of the half-time for the compound CH_2ClCOF shown in Fig. 3 was determined for the 96 peak in the mass spectrum of an unknown mixture

[8] *Handbook of Chemistry and Physics* (Chemical Rubber Publishing Company, Cleveland, Ohio), thirtieth edition, p. 1736.
[9] A. H. Young and S. Dushman, Phys. Rev. **68**, 278 (1945).
[10] Eden, Burr, and Pratt, Anal. Chem. **23**, 1735 (1951).

and served to identify the compound. The identity was later verified by other means. This method of molecular weight determination has proven very useful in this Laboratory. These results are at variance with those of Eden, Burr, and Pratt;[10] however, their leak diameter of 0.043 mm was too large for the pressure they used to fulfill the conditions of pure molecular flow and the deviations they observed were no doubt due to this cause. Benzene, with a mean free path of 0.028 cm at 40μ (their maximum pressure) gives a ratio of 6.5 for mean free path to orifice diameter. This is less than the minimum of 10 recommended by Knudsen. The compounds of higher molecular weight which they used would have an even shorter mean free path.

It might be mentioned at this point that the recommendations made by Honig[1] for the sample pressure and orifice diameter were based on calculations for carbon dioxide, and it was implied that gases with a shorter mean free path would not exhibit molecular flow through this leak at the higher pressures.

The effusion of isotopes would also be expected to occur at different rates. The peaks at mass 50 of CH_3Cl^{35} and 52 of CH_3Cl^{37} in methyl chloride were observed. The time required for the peak heights to decrease to one tenth their initial values were 102.5 minutes for the 50 peak and 104.3 minutes for the 52 peak; which are in the ratio of 1.018; the ratio of the reciprocal of the square roots of the molecular weights is 1.020. In a similar experiment with neon, the ratio of rates for the isotopes of mass 20 and 22 was 1.050, which is again in substantial agreement with the theoretical value of 1.049.

Another consequence of the progressive decrease in pressure is the fact that the fragmentation patterns obtained by scanning will be dependent upon the relative rates of scanning and pressure decrease and on the molecular weight of the compounds. As an example, a typical standard operating procedure for determining a fragmentation pattern[11] might include the rate of leak for *n*-butane −0.25 percent per minute, and a scanning rate of 10 minutes for the range of mass 12 to 58. During that period the pressure will have decreased about 2.5 percent and, hence, the 58 peak is recorded 2.5 percent lower than it would be had a constant pressure been maintained on the leak compared with mass 12 as a standard.

Other peaks will also be lower by an amount depending upon the interval between the introduction of the sample and the time of measurement. The peak height, as a function of time, may be given more precisely in terms of the exponential leak rate. For the leak and sample volume where the half-time was given as $4.6M^{\frac{1}{2}}$, the peak height (p_h) at any time (minutes) is given by

$$p_h = p_{ho} \exp[-0.0132(M)^{\frac{1}{2}}t],$$

where p_{ho} is the peak height if measured at the time the sample is introduced into the spectrometer (t_0). In quantitative analysis with the mass spectrometer, the usual practice[12] is to adopt a rigid time schedule for the introduction of the sample, the time the scanning is started, and the rate of scan. By normalizing all peak heights to zero time, it is not necessary to adopt any schedule. An arbitrary starting time for the calculations is not permissible because of fractionation of the sample, which starts when the sample is first introduced into the sample reservoir. Fragmentation patterns reported on this normalized basis offer a comparison of instrument performance free of the complication of sample pressure variation.

The sensitivity of a compound is generally quoted[11] in scale divisions per micron of pressure on the leak and, for comparison, the sensitivity under comparable conditions with *n*-butane. A more absolute statement of sensitivity would give the deflection in terms of the rate at which the sample passed through the leak. For example, with the spectrometer used, the sensitivity for nitrogen was 14,500 divisions per micron when the flow rate was 0.000230μl/sec. The sensitivity could then be given as 6.3×10^{7} divisions at a flow rate of one micron liter per second, or 1.95×10^{-9} unit of deflection at a flow rate of one molecule per second.

The sensitivity of the spectrometer may also be calculated from these data and the current sensitivity of the amplifier, which in the present case is 5.0×10^{-15} ampere per division, that is, 3.12×10^{4} ions per second will maintain unit deflection. Combining these data, one nitrogen molecule ion (N_2^+) is collected in 1.65×10^{4} molecules that pass through the leak.

CONCLUSION

The necessary conditions for molecular flow through the mass spectrometer leak have been restated in terms of leak diameter, flow rates, and sample pressure in the equations

$$L = 10d = L_0/p,$$
$$Q = 2.86d^2p(T/M)^{\frac{1}{2}}.$$

Experimental verification of molecular flow through the leak was based on the constant ratio of pressure to peak height, and the dependence of the flow rate on the square root of the temperature and on the inverse of the square root of the molecular weight.

It was shown that the laws of viscous flow govern the flow of gas through leaks such as may be detected by the helium leak detector, and as a consequence, argon is a more sensitive probe gas than helium.

The rate of decay of pressure in the sample reservoir affects the fragmentation patterns and sensitivities, and it is suggested that they be reported in some more fundamental terms, such as flow rates, or fractional collection efficiency of the spectrometer.

It is a pleasure to thank Dr. Saul Dushman and Dr. H. A. Liebhafsky for their advice in the preparation of this paper.

[11] American Petroleum Institute Research Project No. 44, *Catalog of Mass Spectra.*

[12] Wasburn, Wiley, and Rock, Ind. Eng. Chem., Anal. Ed. **15**, 541 (1943).

An Integrable Case of Electron Motion in Electric and Magnetic Field

H. Poritsky and R. P. Jerrard
General Engineering Laboratory, General Electric Company, Schenectady, New York
(Received April 10, 1952)

Electron motion is studied in a two-dimensional electric field of potential $V=A+B(x^2-y^2)/2$ and a uniform magnetic field $\mathbf{H}=(0, 0, -H)$ normal to the electric field, where A, B, $-H$ are constants. The equipotential lines of the electric field in any plane $z=$const. consist of rectangular hyperbolas with a point of zero field strength at $x=0$, $y=0$. The differential equations of motion are integrated, and expressions are given for the electron paths. For this type of field, the electron motion consists of superposition of elliptic and hyperbolic motions, that is, of a simple harmonic motion along an ellipse whose center moves along a hyperbola. The latter hyperbolas intersect the equipotential hyperbolas, so that, unlike the uniform crossed field case, the electron may drift into regions of higher or lower potential.

THERE are but few cases where exact, analytic solutions of the equations of electron motion in electric and magnetic fields are available. One such case is furnished by uniform electric and magnetic fields, normal to each other; here the electron describes cycloidal paths (the term cycloidal includes prolate and oblate cycloids) in a plane perpendicular to the magnetic vector, with the axis of the cycloid at right angles to both $\mathbf{E}$ and $\mathbf{H}$. In the following we call attention to another integrable case, namely the case where the magnetic field is uniform, and pointing in the direction of the z axis; the electric field two-dimensional, normal to the magnetic vector, and described by the potential or voltage function

$$\varphi=A+B(x^2-y^2/2), \tag{1}$$

where A and B are constants, $B>0$. In this case the equipotential lines in any plane $z=$const are rectangular hyperbolas, as shown in Fig. 1. At the origin the field vanishes.

The differential equation of motion of the electron in electrostatic and magnetic fields, in vector form, in Gaussian units, is

$$m\ddot{\mathbf{r}}=\epsilon\nabla\varphi-\frac{\epsilon}{c}\dot{\mathbf{r}}\times\mathbf{H}, \tag{2}$$

where the electron charge is represented by $-\epsilon(\epsilon>0)$. Replacing $\mathbf{H}$ by $(0, 0, -H)$,* utilizing Eq. (1), dividing by m and rearranging one obtains the scalar equations

$$\begin{aligned}\ddot{x}-\epsilon H\dot{y}/mc-B\epsilon x/m&=0,\\ \ddot{y}+\epsilon H\dot{x}/mc+B\epsilon y/m&=0,\\ \ddot{z}&=0.\end{aligned} \tag{3}$$

The last equation integrates into

$$\dot{z}=\dot{z}_0, \quad z=z_0+\dot{z}_0 t, \tag{4}$$

where z_0, $\dot{z}_0$ are the initial values of z, $\dot{z}$. We proceed with consideration of the (x, y)-Eqs. (3).

The first two Eqs. (3) are linear in x and y with constant coefficients. A solution of the form

$$\begin{aligned}x&=Le^{\lambda t}\\ y&=Me^{\lambda t},\end{aligned} \tag{5}$$

is assumed where λ, L, M are constant, and substitution of this solution into Eqs. (3) results in

$$\begin{aligned}L(\lambda^2-B\epsilon/m)-\epsilon H\lambda M/cm&=0,\\ +L\epsilon H\lambda/cm+M(\lambda^2+B\epsilon/m)&=0.\end{aligned} \tag{6}$$

These equations admit solutions other than $L=M=0$ when and only when the determinant vanishes:

$$\begin{vmatrix}\lambda^2-B\epsilon/m & -\epsilon H\lambda/mc\\ +\epsilon H\lambda/mc & \lambda^2+B\epsilon/m\end{vmatrix}$$

$$=\lambda^4+(\epsilon H/mc)^2\lambda^2-(B\epsilon/m)^2=0. \tag{7}$$

For any root of Eq. (7) the Eqs. (6) become linearly dependent and yield

$$L/M=-\epsilon H\lambda/c[-m\lambda^2+B\epsilon]=c[m\lambda^2+B\epsilon]/-\epsilon H\lambda. \tag{8}$$

Equation (7) is quadratic in λ^2. To obtain its roots in

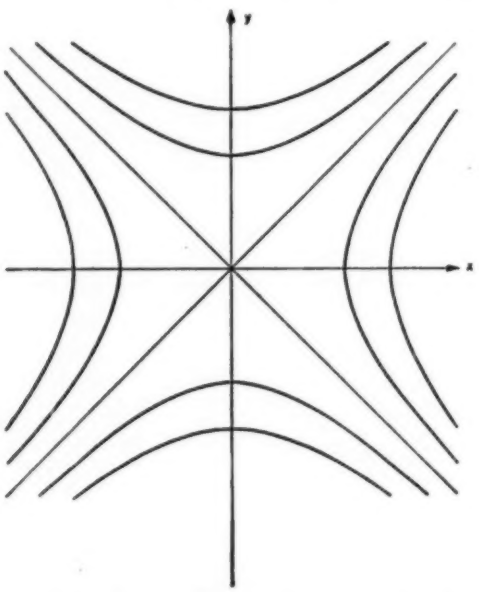


Fig. 1. Equipotentials in the electric field of Eq. (1).

* With H a positive number the field points in the direction of the negative z axis; the electrons will then circle in a clockwise direction.

dimensionless form we rewrite it as

$$\lambda^4+\mu^2\lambda^2-\alpha^2\mu^4=0, \tag{9}$$

where

$$\mu=\epsilon H/mc, \qquad \alpha=Bc^2m/H^2\epsilon=B\epsilon/m\mu^2. \tag{10}$$

Here μ is the Larmor precession frequency, that is, the rotational frequency of an electron in the magnetic field H (in absence of any electric field), and α is a dimensionless parameter involving the electric and magnetic field.† The solution of Eq. (7) yields four values for λ, two real, two pure imaginary:

$$\lambda=\pm\mu[-1+(1+4\alpha^2)^{\frac{1}{2}}/2]^{\frac{1}{2}}=\pm\eta, \qquad \lambda=\pm i\mu[1+(1+4\alpha^2)^{\frac{1}{2}}/2]^{\frac{1}{2}}=\pm i\omega. \tag{11}$$

The ratio M/L may also be written using the parameters μ, α, thus

$$-M/L=(\alpha\mu^2-\lambda^2)/\mu\lambda=\alpha\mu/\lambda-\lambda/\mu. \tag{12}$$

This ratio also has four values, one corresponding to each value of λ in Eq. (11) and will be written

$$M/L=\pm[-2a+(1+4\alpha^2)^{\frac{1}{2}}]^{\frac{1}{2}}=\pm g, \qquad M/L=\pm i[2a+(1+4\alpha^2)^{\frac{1}{2}}]^{\frac{1}{2}}=\pm ih. \tag{13}$$

For the real λ's Eq. (5) becomes

$$x=Le^{\pm\eta t}, \quad y=\pm gLe^{\pm\eta t}, \tag{14}$$

where L is a constant and represent motions along the straight line

$$y=+gx, \tag{15}$$

receding from the origin, and along the straight line

$$y=-gx, \tag{16}$$

approaching the origin.

Combining linearly the four special solutions of Eq. (11) one obtains the following general (real) solutions:

$$x=S\cosh(\eta t+\gamma)+T\cos(\omega t+\beta), \qquad y=-gS\sinh(\eta t+\gamma)-hT\sin(\omega t+\beta), \tag{17}$$

and

$$x=-S/g\sinh(\eta t+\gamma)+T/h\sin(\omega t+\beta), \qquad y=S\cosh(\eta t+\gamma)+T\cos(\omega t+\beta), \tag{18}$$

where S, T, γ, β are arbitrary constants.

For $T=0$ there results a motion along a hyperbola of transverse semi-axis S whose asymptotes are given by Eqs. (15), (16); Eqs. (17) yield a motion along hyperbolas crossing the x axis; Eqs. (18) along hyperbolas crossing the y axis. For $S=0$, the solutions (17), (18) represent a simple harmonic motion of frequency ω, exceeding μ, which is time quadrature and describes an ellipse

$$x^2/a^2+y^2/h^2a^2=1 \tag{19}$$

† One may interpret α as R^2/x^2 where R is the radius of the circular electron motion with velocity v in the magnetic field H, and x is the distance of fall of an electron along the x axis in the field of Eq. (1), starting from rest at the origin until it has acquired the (same) velocity v.

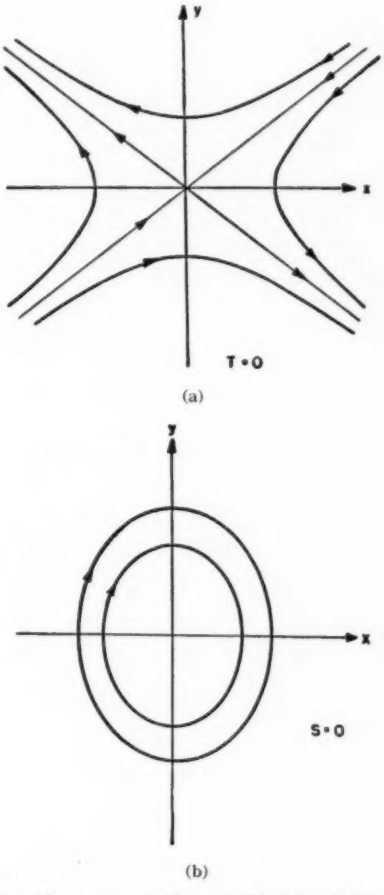


FIG. 2. Schematic representation of elliptic and hyperbolic component motions of the electron.

in clockwise fashion. These special motions are shown schematically in Fig. 2.

The general solution (17), (18) which have been obtained above thus represents a motion which can be described as a superposition of a simple harmonic motion in the x and y directions of frequency ω, with amplitudes whose ratio is h, and a motion of the center point about which this oscillation takes place along a hyperbola.

Further solutions which are not contained in Eqs. (17), (18) are obtained by using the exponential solutions (15) with the T-terms and correspond to motions along ellipses with the center point moving along the asymptotes (16), (17) rather than along the hyperbola proper.

The above solutions satisfy the energy equation which is valid for motion in a magnetic field:

$$(\dot{x}^2+\dot{y}^2)-\frac{2\epsilon}{m}\varphi=\text{const}, \tag{20}$$

though this has not been utilized in obtaining them.

In Eqs. (17) and (18) there are four constants, S, T, γ, β which are determined by the conditions of initial

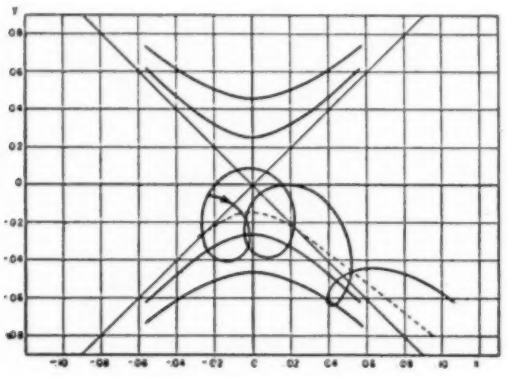


FIG. 3. Electron orbit.

position and velocity of the electron. Thus, if

$$\begin{aligned} x&=x_0,\\ y&=y_0,\\ &\text{for}\quad t=0\\ dx/dt&=\dot{x}_0,\\ dy/dt&=\dot{y}_0, \end{aligned} \tag{21}$$

solution of the four equations which result by applying (21) to the solution (17) yields for the values of the constants γ, S, β, T

$$\begin{aligned} \tanh\gamma&=(y_0\omega+\dot{x}_0h)/(x_0\omega h-\dot{y}_0),\\ S\cosh\gamma&=(\dot{y}_0-\omega hx_0)/(\eta g-\omega h),\\ \tan\beta&=(g\dot{x}_0-y_0\eta)/(x_0\eta g-\dot{y}_0),\\ T\cos\beta&=(x_0\eta g-\dot{y}_0)/(\eta g-\omega h). \end{aligned} \tag{22}$$

Similarly, if Eqs. (18) are used, there results for the determination of the constants

$$\begin{aligned} \tanh\gamma&=(x_0\omega h-\dot{y}_0)/(y_0\omega+\dot{x}_0h),\\ S\cosh\gamma&=(\dot{x}_0+\omega gy_0)/(\eta h+\omega g),\\ \tan\beta&=(x_0g\eta-\dot{y}_0)/(y_0\eta-\dot{x}_0g),\\ T\cos\beta&=(y_0\eta h-\dot{x}_0)/(\eta h+\omega g). \end{aligned} \tag{23}$$

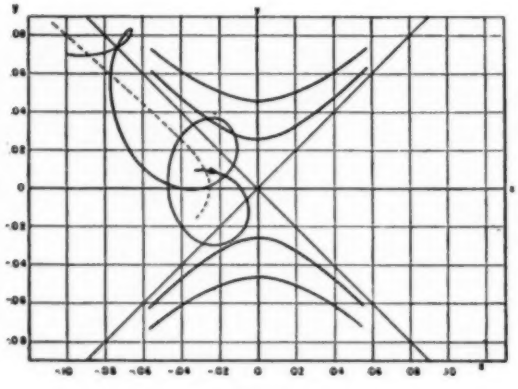


FIG. 4. Electron orbit.

Whether the set of Eqs. (17) or (18) should be used for the electron path corresponding to the initial conditions (21) is determined by calculating $\tanh\gamma$. Note that the two values of $\tanh\gamma$ given above are the reciprocals of each other. Since $|\tanh\gamma|$ can never be greater than unity, this tells us which of the two solutions must be used for any particular path.

If $|\tanh\gamma|=1$, then neither solution (17), (18) applies, and the solution reduces to a combination of the T-terms with the solutions (15).

Comparing the above motion with the motion of an electron in a uniform electric field, the following differences are evident. In the uniform field the motion consists of a circular motion with Larmor frequency μ combined with a rectilinear motion parallel to the equipotentials. In the present field the circular motion is replaced by an elliptic motion of frequency $\omega > \mu$, while the motion along the equipotential is now replaced by

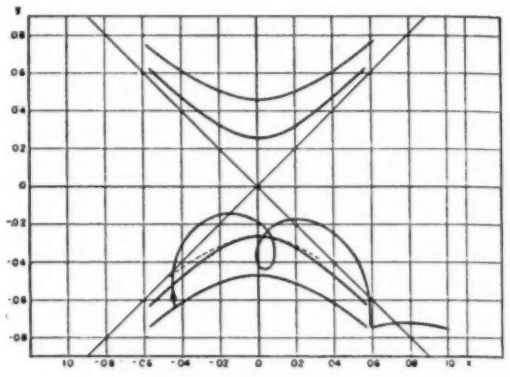


FIG. 5. Electron orbit.

the motion along a hyperbola with asymptotes given by Eq. (16). Note that these hyperbolas cross the equipotential lines of the field, and allow the electron to cross into regions of higher or lower potential, so the electrons have a chance to fall down or rise up the potential hill.

Examples of motion in this type of field are shown on Figs. 3, 4, and 5. The dotted lines are the hyperbolas of Fig. 2 about which the elliptical oscillatory motion of the electron takes place. The solid hyperbolas are equipotential lines.

The above solution may be used to furnish a first order correction for electron motions in general two-dimensional electrostatic (space-charge free) fields and constant magnetic field by expanding the complex field potential as a Taylor series and identifying the linear and quadratic terms with the potential (1) for proper A, B and proper positions of the origin and direction of the real axis.

Letters to the Editor

Possible Explanation of Briggs' Limiting Negative Pressure Data

V. J. BERRY AND E. W. HOUGH
Stanolind Oil and Gas Company, Tulsa, Oklahoma
(Received March 27, 1952)

BRIGGS[1,2] has recently published experimental results for the limiting negative pressures of water, aniline, benzene, and several other materials as functions of temperature. During the experiments the liquid was confined in a glass capillary tube, and it was noted by Briggs that the observed rupture of the liquid under tension may have actually been initiated at the glass-liquid interface, rather than within the body of the liquid itself. Any proposed mechanism must explain the dependence of the limiting pressure upon temperature.

Furth[3] has developed from the hole theory of liquids an equation relating the maximum attainable supersaturation pressure to surface tension and temperature. Furth's equation is

$$p_0 - p^* = B\sigma^4(kT)^{-3}, \qquad (1)$$

where p_0 is the vapor pressure at absolute temperature T, p^* is the minimum prevailing pressure attainable prior to the appearance of a gas phase, σ is the surface tension at the prevailing pressure and temperature, and k is Boltzmann's constant. B is a dimensionless coefficient which Furth evaluates by two different approximate methods as 1.3 and 1.57, respectively.

In Fig. 1 are shown the data of Briggs for aniline, benzene, and water. Some data near the freezing points have been omitted in the case of aniline and benzene. The curves represent theoretical values of $p_0 - p^*$ calculated from Eq. (1), using experimental values for the surface tensions at atmospheric pressure as a function of temperature.[4] A single value of 0.198 for the constant B was used in the theoretical calculations for both aniline and benzene, whereas the value of B chosen for water was 0.085.

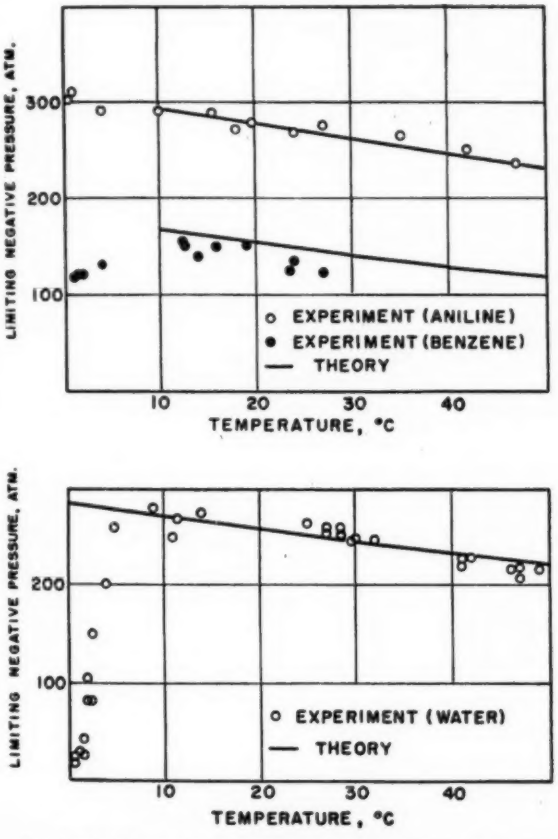


FIG. 1. Limiting negative pressures of aniline, benzene, and water (see references 1 and 2).

In view of the good agreement in reproducing theoretically the observed temperature dependence it seems reasonable to assume that the mechanism of bubble formation being observed above 0° for those materials is that of hole formation in the bulk of the liquid. The fact that a single value of the coefficient B may be used to correlate the data for both benzene and aniline lends credence to this hypothesis. The difference between the theoretical curves for these two substances may be a result of differences in their surface tensions. The disagreement between the experimental and the theoretical values of the dimensionless coefficient B may be attributable to uncertainties in the absolute values of surface tension at large negative pressures and to approximations involved in Furth's analysis.

The theory already mentioned[3] is powerless to explain the observed decrease in limiting pressures as the temperature is lowered to the freezing point. We suggest that this anomaly may be a result of surface effects at the liquid-glass interface becoming important near the freezing point. Although Furth[5] and Fisher[6] developed theories which express the effect of contact angle on supersaturation attainable, practically no information on the variation of contact angle with temperature is available in the literature. Limiting negative pressure experiments of the Briggs type performed with known surface effects would be of interest. The contact angle for several members of a homologous series would have to be established experimentally as a function of temperature near the freezing point, and two adajacent members having zero contact angle and a small, nonzero contact angle, respectively, could then be chosen. For instance, according to Fox and Zisman,[7] at 20°C, pentane has a contact angle of 0 deg on polytetrafluoroethylene, while heptane has a contact angle of 21 deg. Polytetrafluoroethylene is suggested for the solid material because the contact angles obtained with this material are reproducible, and the spreading pressure is very low for a large variety of fluids.[7] The effect of this latter property would aid materially in evaluating the influence of the solid-liquid interface on any anomaly that might be observed.

[1] L. J. Briggs, J. Appl. Phys. **21**, 721 (1950).
[2] L. J. Briggs, J. Chem. Phys. **19**, 970 (1951).
[3] R. Furth, Proc. Cambridge Phil. Soc. **37**, 252 (1941).
[4] International Critical Tables **4**, 454, 447.
[5] R. Furth, Proc. Cambridge Phil. Soc. **37**, 276 (1941).
[6] J. C. Fisher, J. Appl. Phys. **19**, 1062 (1948).
[7] H. W. Fox and W. A. Zisman, J. Colloid Sci. **5**, 514 (1950).

Temperature Gradient Owing to Radiation Absorption

G. B. SHOOK
Michelson Laboratory, U. S. Naval Ordnance Test Station, Inyokern, China Lake, California
(Received March 17, 1952)

EXPERIMENTAL and mathematical studies have been made of the temperature distribution $U(x, t)$ in a medium absorbing a time-dependent flux of radiation. The mathematical treatment is based on the following set of assumptions:

(1) Absorbing medium consists of a homogeneous semi-infinite region $x>0$. (2) Radiant flux $I(t)$ is in x direction only, and reflection from surface $x=0$ is ignored. (3) The constants K, c, ρ (thermal conductivity, specific heat, and density, respectively) are independent of temperature. (4) Reradiation is from surface only, into surroundings at a constant temperature U_a and at a rate proportional to the temperature difference $U_{x=0} - U_a$. The equation to be solved is then

$$\partial U/\partial t = \kappa \partial^2 U/\partial x^2 + I(t)\alpha/c\rho \exp(-\alpha x), \qquad (1)$$

subject to the following boundary conditions: (a) $(\partial U/\partial x)_{x=0}$

$=h(U_{x=0}-U_a)$, (b) $\lim_{x\to\infty}(\partial U/\partial x)=0$, (c) $U=0$ when $t=0$; where $\kappa=K/c\rho$, α is the average absorption coefficient of the medium, h is the rate constant for the combined processes of reradiation and conduction from the surface to the surroundings.

Equation (1) is then converted to an ordinary differential equation by means of the Laplace transformation with the solution[1]

$$\bar{U}=[hU_a/(h+q)s]\exp(-qx)+\varphi(s)\{\exp(-\alpha x)-[(h+\alpha)/(h+q)]\exp(-qx)\}/(q^2-\alpha^2), \quad (2)$$

where $\bar{U}$ is the transformed temperature variable, $q^2=s/\kappa$, s is the transform parameter, and $\varphi(s)$ is the transform of $I(t)$ multiplied by α/K.

The manufacturer's data on commercial photoflash bulb types *SM* and No. 5 give a curve of lumens *versus* time, which, when converted to energy/cm² *versus* time by means of the standard visibility curve and conversion factors, corresponds rather well to the empirical function $I(t)=at^4\exp(-bt)$.

Because of experimental interest in the results, the above

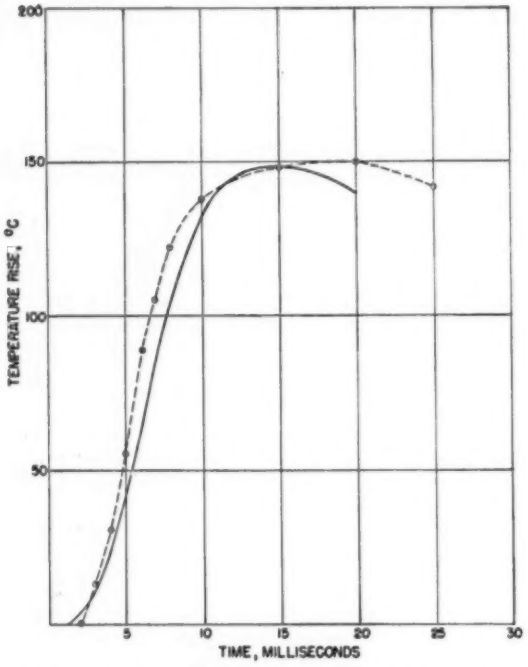


FIG. 1. Comparison of calculated (——) and observed (—oo—) temperature rise at face of sample 1.9-cm from center of photoflash bulb type *SM*. Constants for plastic sample used in Eq. (3) are: $\alpha=126$ cm^{-1}, $\kappa=8.4\times10^{-4}$ cm²/sec, $a=5.25\times10^{12}$ $c=0.37$ cal/g°C, $\rho=1.6$ g/cm³, $b=667$.

empirical function was used to obtain $\varphi(s)$. The inversion of Eq. (2) then gives the following result:

$$U=-U_a[e^{\lambda x+\lambda^2\kappa t}\operatorname{erfc}(h(\kappa t)^{\frac{1}{2}}+\tfrac{1}{2}x(\kappa t)^{-\frac{1}{2}})-\operatorname{erfc}\tfrac{1}{2}x(\kappa t)^{-\frac{1}{2}}]$$
$$+(\alpha a/c\rho)\Big\{e^{\alpha^2\kappa t-\alpha x}\int_0^t\tau^4e^{-\beta\tau}d\tau$$
$$+(h+\alpha)\int_0^t[\kappa he^{xh+h^2\kappa\tau}\operatorname{erfc}(h(\kappa\tau)^{\frac{1}{2}}+\tfrac{1}{2}x(\kappa\tau)^{-\frac{1}{2}})$$
$$-\kappa(\pi\tau)^{-\frac{1}{2}}\exp(-x^2/4\kappa\tau)d\tau]\cdot\int_0^{t-\tau}\lambda^4e^{-\beta\lambda}d\lambda\Big\}, \quad (3)$$

where $\beta=b+\alpha^2\kappa$.

Measurements have been made of the surface temperature rise of a plastic material exposed to the flash of the two bulb types mentioned. The technique involved placing a 0.001-in. diameter nickel wire on the surface of the plastic in such a manner that the upper wire surface was tangent to the sample surface, the body of the wire being embedded in the plastic. This wire was then used as a resistance thermometer. A comparison of calculated *versus* observed temperature rise is shown in Fig. 1. The calculated temperatures were obtained from Eq. (3) assuming that $x=h=0$.

[1] H. S. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids* (Oxford University Press, London, 1947), first edition, p. 249, give solution of a related problem where $I(t)=kt^{n/2}$.

Single Synthetic Cadmium Sulfide Crystals

S. J. CZYZAK,* D. J. CRAIG,† AND C. E. MCCAIN
University of Detroit, Detroit, Michigan
AND
D. C. REYNOLDS, *Battelle Memorial Institute, Columbus, Ohio*
(Received May 9, 1952)

SINGLE crystals of cadmium sulfide have been grown by Frerichs and Lorenz[1,2] by the reaction of cadmium vapor in hydrogen sulfide. Cadmium metal was used in this reaction. The authors have grown single crystals of cadmium sulfide by a method similar to that used by Reynolds and Czyzak.[3]

The method of growing these crystals is as follows: Chemically pure cadmium sulfide powder was distributed in the center of a quartz tube over a length of 2 to 3 inches. After inserting the quartz tube in a combustion furnace, the tube was evacuated and hydrogen sulfide gas was passed into the system until a pressure of 6 psi was obtained. The tube and its contents were then heated to a temperature of 1000°C for a period of 48 to 72 hours. The crystals grew in the form of hexagonal prisms. Angles of 30° were obtained when crystals were optically measured between neighboring prism faces. The simplest interpretation would seem to be that the prism faces are planes of two types, $11\bar{2}0$ and $10\bar{1}0$, with the *C* axis in the direction of the edge of the dihedral angle between any of these two planes. The average size of the crystals obtained by this method was approximately 3×3×6 mm, with the larger ones reaching a size of approximately 3×5×10 mm. This is approximately the same size that was obtained in growing zinc sulfide crystals using this method. For best results, it was found that crystals grown by this method should be cooled at a rate of about 10°C per hour. With crystals grown by this method, one does not obtain the ribbon and needle-like crystals that one gets when one uses the method of Frerichs. The optical absorptions of the cadmium sulfide crystals was measured with a photospectrometer at room temperature to determine the absorption cutoff. The data obtained are shown in Fig. 1. The absorption cutoff

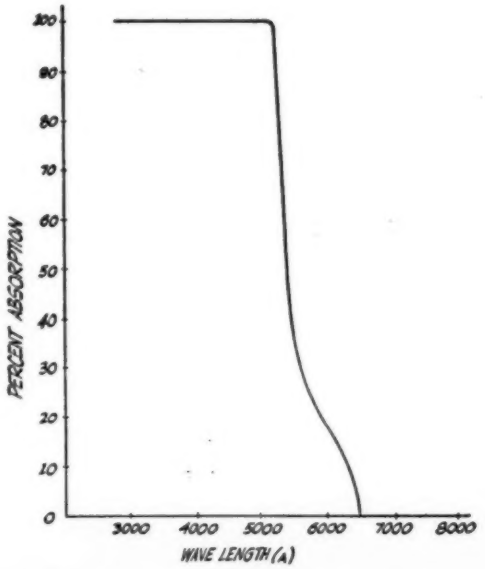


FIG. 1. Optical absorption of cadmium sulfide.

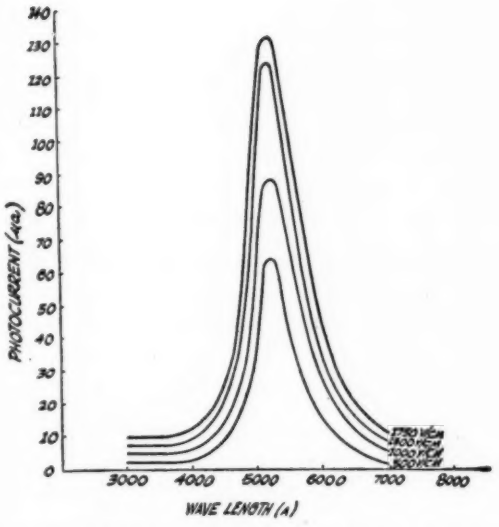


FIG. 2. Spectral distribution of photoelectric sensitivity of cadmium sulfide for various field intensities.

for the cadmium sulfide crystals was found to be at 5200A, and they are entirely opaque below this value. This agrees with the value obtained by Frerichs and others. The absorption decreases until, in the red region, the crystals are completely transparent. However, the curve at the high wavelength differs from that of Frerichs in that Frerichs obtains complete transmission at approximately 5900A, whereas we obtain complete transmission at 6500A. The long wavelength tail of the absorption curve varies among crystals indicating that foreign impurities or defects may be more pronounced in some crystals than in others.

The purity of the crystals can be determined to some extent by their luminescent properties. Some of the crystals showed a red fluorescence when exposed to ultraviolet light while the majority of the crystals showed no evidence of luminescence. Crystals showing no signs of luminescence are considered to be quite pure. No effort was made to correlate the long wavelength tail of the absorption curve with luminescence in the crystals. In general, both zinc sulfide and cadmium sulfide crystals grown in a hydrogen sulfide atmosphere can be produced in which neither phosphorescence no fluorescence exists. However, in any batch that is grown, some of the crystals produced will fluoresce or phosphoresce. Crystals which did not fluoresce or phosphoresce were used in our tests.

The single crystals of CdS were also tested for photoconductivity. The photoconductivity of the cadmium sulfide crystals was measured as a function of wavelength. Figure 2 gives the spectral distribution of the photocurrent in microamperes for various field intensities. The maximum photoelectric current is at 5200A, which is also the absorption cutoff for the CdS crystal.

* Now Lt. Col., USAF, Flight Research Laboratory, WADC, Wright-Patterson Air Force Base, Ohio.
† Now Research Staff Member, Wyandotte Chemicals Company, Wyandotte, Michigan.

[1] R. Lorenz, Chem. Ber. **24**, 1509 (1891).
[2] R. Frerichs, Phys. Rev. **72**, 594 (1947).
[3] D. C. Reynolds and S. J. Czyzak, Phys. Rev. **79**, 543 (1950).

Announcements

ONR Sponsored Conference on Magnetism

The Office of Naval Research is sponsoring a Conference on Magnetism to be held at the University of Maryland from September 2 to 6, 1952. Several foreign visitors, among them Professors Neel, Stoner, Bates, Sucksmith, and Gorter, are expected to attend and present papers. The conference will cover the following broad subjects: Technical Magnetization, Exchange Forces, Antiferromagnetism, Ferrites, Experimental Techniques, and Paramagnetism. An evening session devoted to a round table discussion of exchange phenomena and theories of the origin of ferromagnetism with the participation of Professors Neel, Slater, Stoner, Van Vleck, and Zener is also projected as a part of the program. The conference is being organized by the ONR magnetism panel whose Chairman is Professor R. Smoluchowski. The program committee consists of R. M. Bozorth, J. E. Goldman, G. T. Rado, and J. H. Van Vleck, Chairman. A local committee has been organized in Washington, D. C., to arrange for the meeting comprising E. Montroll, L. R. Maxwell, L. Marton, and G. T. Rado, Chairman. Inquiries concerning attendance and housing should be directed to Dr. Rado at the Naval Research Laboratory, Washington, D. C. It is expected that the foreign visitors will spend some time in this country visiting various centers of magnetic research. Communications concerning the scheduling of their visits to laboratories or universities should be addressed to Dr. V. Wilson, General Electric Research Laboratory, Schenectady, New York.

Visiting Lecturers at Harvard University

The Division of Applied Science of Harvard University announces the following visiting lecturers for the fall term of the academic year 1952–1953: *Dr. Walter H. Brattain* of the Bell Telephone Laboratories, Murray Hill, New Jersey; *Dr. Sydney Goldstein*, Vice President of The College of Technology in Haifa, Israel; and *Dr. W. Duncan Rannie*, Associate Professor of Mechanical Engineering at California Institute of Technology.

Dr. Walter Brattain is one of the co-inventors of the original transistor. The invention occurred as a result of research work done at the Bell Telephone Laboratories. The transistor is a device that can perform many of the functions of the vacuum tube and that is aptly described in the March issue of *Harpers* as "a little gadget with a large future." While at Harvard Dr. Brattain will give with the assistance of two instructors in Cruft Laboratory a course connected with the use of transistors in electrical circuits. In addition, he will conduct jointly with Professor Harvey Brooks of the Division of Applied Science a seminar course on the questions of solid state physics underlying the behavior of the transistor. Harvard is one of the first two universities to have two courses of a semester's length on the subject of transistors. The other is the University of Illinois, where Professor John Bardeen, the other co-inventor of the transistor and formerly of the Bell Telephone Laboratories, is now located.

Professor Goldstein has an international reputation in the field of fluid mechanics. He was for many years Professor of Applied Mathematics at Manchester University, but because of his strong nationalist feeling he left this post in 1950 to take a position in Israel. He is at present Vice President of the College of Technology in Haifa. He is also a Fellow of the Royal Society and a foreign member of the Royal Netherlands Academy of Science and Letters. Professor Goldstein will give an advanced course in hydrodynamics and aerodynamics.

Professor Rannie will give two half-courses on the theory of heat transfer, one on conduction and radiation and the other on convection. Professor Rannie is a member of the Subcommittee on Compressors and Turbines of the NACA. He is also a member of the Scientific Advisory Board to the Air Forces.

National Science Foundation Appoints Ralph A. Morgen to Direct Engineering Research Support Program

Dr. Ralph A. Morgen, Director of the Engineering and Industrial Experiment Station of the University of Florida, has been appointed to the staff of the Division of Mathematical, Physical, and Engineering Sciences of the National Science Foundation. Dr. Morgen will be in charge of the Foundation's research support program in the engineering sciences.

Conference on Gaseous Electronics

The fifth Conference on Gaseous Electronics, sponsored by the Division of Electron Physics of the American Physical Society, will be held at Princeton University and at the RCA Laboratories in Princeton, New Jersey, on September 4, 5, and 6, 1952. The program will consist of invited and contributed papers pertaining to the fundamental physics of gas discharge phenomena. The technical sessions will be held in Frick Hall of Princeton University and at RCA Laboratories.

1953–1954 Fulbright Awards

Announcement is being made for the 1953–1954 competition for Fulbright awards for university lecturing and postdoctoral level research in *Europe and the Near East*. Included in this competition are awards for Australia, Belgium and Luxembourg, Denmark, Egypt, France, Greece, Iraq, Italy, Japan, Netherlands, Norway, Pakistan, Turkey, Union of South Africa, United Kingdom, and Colonial Dependencies. A Fulbright Agreement has recently been signed with the Union of South Africa and a very small program will be initiated for 1953–1954.

The Committee on International Exchange of Persons is preparing a detailed program booklet concerning the awards. The booklets are expected to be available early in July. The *closing date for making application is October 15, 1952.*

Conference on Instruments and Measurements in Stockholm

A Conference on Instruments and Measurements, sponsored by the Swedish Academy of Engineering Sciences and the Association of Technical Physicists, will be held in Stockholm, Sweden, September 23-30, 1952. An exhibition under the same sponsorship will be held concurrently. Fifty-three speakers will represent 10 foreign countries, as follows: Denmark, 1; England, 10; France, 1; Germany, 10; Holland, 4; Norway, 1; Sweden, 15; Switzerland, 4; U. S. A., 6. Proceedings will be published and available at $6.00 each. The Secretary of the Conference is Mr. Yngve Axner, Kingl. Tekniska Hogskolan, Stockholm.